

Sequence Name	Description	Length	Init. Score	Opt. Score	Sig. Frame
---------------	-------------	--------	-------------	------------	------------

X	10	20	30	40	50	60	70
TCAAGSAAACCGCTACGACCTGGCCCGGGTGGCCGCTACAAAGTCAACAAAGAAACCTGGGCTAACACCGC							
X	10	20	30	40	50	60	70
TCAAGSAAACCGCTACGACCTGGCCCGGGTGGCCGCTACAAAGTCAACAAAGAAACCTGGGCTAACACCGC							
X	10	20	30	40	50	60	70
TCAAGSAAACCGCTACGACCTGGCCCGGGTGGCCGCTACAAAGTCAACAAAGAAACCTGGGCTAACACCGC							





	150	160	170	180	190	200	X
ACGAGGGGACACCGCAGTGAAGTCCCGCGTGGGGCTGCAGGTGCGGGGTGGAGACCCGACCACTT							
ATGAGGGTACTGCACGATGACCGTTCCAGGTGGGGTCCAGGTCCGATGATCTACTGACGACACTT							
150	160	170	180	190	200	X	

[illegible]

14. US-09-697-123B-24 (1-208)  
US-09-697-123B-6 Sequence 6, Application US/09697123B

X	10	20	30	40	50	60	70
TCAAGGAGAAGCCCTACGACCTGGGCCGGGTGGCGCATCAAGGTCAACAAGAACTCGGGCTGAACACC							
TCAAAGGAGAAGCCCTACGACCTGGGCCGGGTGGCGCATCAAGGTCAACAAGAACTCGGGCTGAACACC							
TCAAAGGAGAAGCCCTACGACCTGGGCCGGGTGGCGCATCAAGGTCAACAAGAACTCGGGCTGAACACC							
TCAAAGGAGAAGCCCTACGACCTGGGCCGGGTGGCGCATCAAGGTCAACAAGAACTCGGGCTGAACACC							
X	10	20	30	40	50	60	70

AGAATGCCCAACGACAGCAGCCGTGCACCGAAGAGACCTGTCGCACCATGAATACCTGTCGGCTTGG  
- - - - -  
CGGAGCCCATCACGTGTGTCGACTGCAGCCAAGAACCTCTGGCCACATCAAAATATGTCGGCTTG

80 90 100 110 120 130 140

150 160 170 180 190 200 X  
ACGAGGCGACCGACGATGATGAAGTCCCGCGGTGGAGTGTCCCGTGGAGACCGACACAT  
ACGAGGCGTACAGACACGAGATGACCGTCCGCGGGGGGTGAGATGCGCGGTGAACCCACACAT  
150 160 170 180 190 200 X

15. US-09-697-123B-24 (1-208)  
US-09-697-123B-5 Sequence 5, Application US/09697123B

Initial Score	=	172	Optimized Score	=	172	Significance	=	0.51
Residue Identity	=	82%	Matches	=	172	Mismatches	=	36
Gaps	=	0	Conservative Substitutions	=			=	0

X 10 20 30 40 50 60 70  
 TCACAGAGAAGCGCTACGACCTGGCCGGGTGGCCGTACACAGTCAACAAGAACTCGGCTGACACCGC  
 TCACAGAGAAGCGCTACGACCTGGCCGGGTGGCCGTACACAGTCAACAAGAACTCGGCTGACACCGC  
 TCACAGAGAAGCGCTACGACCTGGCCGGGTGGCCGTACACAGTCAACAAGAACTCGGCTGACATGTC  
 X 10 20 30 40 50 60 70

AGATGCGCAACCAACCAGACCCGTGCAGAGAGACTCTGCCACCATCATAAATCTGGTGGCGCTTG  
| | | | |  
GCGAGCCATTCACGTCTGTGACGCTGCAGAACAGACACTGTGGCCACCATCAATATTTGGTCCGCTTGG

80 90 100 110 120 130 140  
80 90 100 110 120 130 140

150 160 170 180 190 200  
 ATGAGGGGACCGCCACGATGAAGTCCCGTGGCGTGGAGTGCCGGTGGAGACCGACGACAT  
 TTTTGGTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT  
 150 160 170 180 190 200  
 ACAGGGCTGAGACACAGATGACCGTTCCGGGGCGGTGAGGTGGCGGTGAACCGACGACAT

16. US-09-697-123B-24 (1-208)  
US-09-697-123B-2 Sequence 20, Application US/09697123B

Initial Score	Optimized Score	Significance
155	166	0.15
Residue Identity = 82%	Matches = 172	Mismatches = 33
Conservative Substitutions = 3		

X 10 20 30 40 50 60  
 TCAAGAGAGAGCGCTACGACCTGCGCGGTGGCGCGCTCAAGGTAAACAGAACTGGGCTGAAC  
 X 10 20 30 40 50 60  
 TCAAGAGAGAAACGCTACGACCTGCGCGGTGGCGCGCTCAAGGTAAACAGAACTGGGCTGAAC  
 X 10 20 30 40 50 60  
 TCAAGAGAGAAACGCTACGACCTGCGCGGTGGCGCGCTCAAGGTAAACAGAACTGGGCTGAAC

[illegible]

ACGAGGGGCACGCCACATCAAGTCCCCCGGTGGCTCGAGGTGCGGGTGGAAGCCGACGACAAT  
150 160 170 180 190 200 X

|||||||  
ACGAGGGGCACGCCACATCAAGTCCCC--GGCATCAGAGGTGCGGGTGGAAGCCGACGACAAT  
150 160 170 180 190 200 X

17. US-09-697-123B-24 (1-208)  
US-09-697-123B-1 Sequence 12, Application US/09697123B

Initial Score	=	144	Optimized Score	=	167	Significance	=	-0.08
Residue Identity	=	82%	Matches	=	172	Mismatches	=	35
Gaps	=	1	Conservative Substitutions	=			=	0

[illegible]

AGAAATGGCCACCAACGACACCTCTGCAAGAGGACCTGTCGCACATGCAATCTGTGGCTTGG  
80 90 100 110 120 130 140  
GCGAGCGATCACACGTTGCGACGCTGGACCGAAGAGATGCTCCCAACCATGACGACTGGTTCGGTGGC  
80 90 100 110 120 130 140

150 160 170 180 190 200  
 ACAGGGGGCAGCCGACGATGAGAGTCCCCGGTGGAGTGGCGGTGAGACCGACACAT  
 |||||  
 150 160 170 180 190 200  
 ACAGGGCCGACGACGATGACCGTTCCCGGGGACCGAGTGGCGGTGAGACCGACACAT

18. US-09-697-123B-24 (1-208)  
US-09-697-123B-1 Sequence 13, Application US/09697123B

Initial Score	=	140	Optimized Score	=	168	Significance	=	-0.17
Residue Identity	=	79%	Matches	=	178	Mismatches	=	30
Gaps	=	15	Conservative Substitutions	=	0			

X 10 20 30 40 50 60 70  
TCAGAGAGAAAGCGCTACGACCTGGCCCGGGTGGCCGCTACAAAGGTCAAAGAACTGGGGCTAACACCG  
TCAGAGAGAGAAAGCGCTACGACCTGGCCCGGGTGGCCGCTACAAAGGTCAAAGAACTGGGGCTAACACCG  
TCAGAGAGAGAAAGCGCTACGACCTGGCCCGGGTGGCCGCTACAAAGGTCAAAGAACTGGGGCTAACACCG  
X 10 20 30 40 50 60 70

**AGATTCGGCAACCA**CACCACCCCTTACCGAAGSAGCCTGTCGCCACCATTGAATTACTGTGCCTTGCC  
| | | |  
**AATACCATTAACA**CCACCAACCGTTACCGAAAGAAGCACCGTAGTCTGCTGGTGGCCTGCTGC  
80      90      100      110      120      130      140

150  
 160  
 170  
 180  
 190  
 200  
 210  
 220  
 230  
 240  
 250  
 260  
 270  
 280  
 290  
 300  
 310  
 320  
 330  
 340  
 350  
 360  
 370  
 380  
 390  
 400  
 410  
 420  
 430  
 440  
 450  
 460  
 470  
 480  
 490  
 500  
 510  
 520  
 530  
 540  
 550  
 560  
 570  
 580  
 590  
 600  
 610  
 620  
 630  
 640  
 650  
 660  
 670  
 680  
 690  
 700  
 710  
 720  
 730  
 740  
 750  
 760  
 770  
 780  
 790  
 800  
 810  
 820  
 830  
 840  
 850  
 860  
 870  
 880  
 890  
 900  
 910  
 920  
 930  
 940  
 950  
 960  
 970  
 980  
 990  
 1000  
 1010  
 1020  
 1030  
 1040  
 1050  
 1060  
 1070  
 1080  
 1090  
 1100  
 1110  
 1120  
 1130  
 1140  
 1150  
 1160  
 1170  
 1180  
 1190  
 1200  
 1210  
 1220  
 1230  
 1240  
 1250  
 1260  
 1270  
 1280  
 1290  
 1300  
 1310  
 1320  
 1330  
 1340  
 1350  
 1360  
 1370  
 1380  
 1390  
 1400  
 1410  
 1420  
 1430  
 1440  
 1450  
 1460  
 1470  
 1480  
 1490  
 1500  
 1510  
 1520  
 1530  
 1540  
 1550  
 1560  
 1570  
 1580  
 1590  
 1600  
 1610  
 1620  
 1630  
 1640  
 1650  
 1660  
 1670  
 1680  
 1690  
 1700  
 1710  
 1720  
 1730  
 1740  
 1750  
 1760  
 1770  
 1780  
 1790  
 1800  
 1810  
 1820  
 1830  
 1840  
 1850  
 1860  
 1870  
 1880  
 1890  
 1900  
 1910  
 1920  
 1930  
 1940  
 1950  
 1960  
 1970  
 1980  
 1990  
 2000  
 2010  
 2020  
 2030  
 2040  
 2050  
 2060  
 2070  
 2080  
 2090  
 2100  
 2110  
 2120  
 2130  
 2140  
 2150  
 2160  
 2170  
 2180  
 2190  
 2200  
 2210  
 2220  
 2230  
 2240  
 2250  
 2260  
 2270  
 2280  
 2290  
 2300  
 2310  
 2320  
 2330  
 2340  
 2350  
 2360  
 2370  
 2380  
 2390  
 2400  
 2410  
 2420  
 2430  
 2440  
 2450  
 2460  
 2470  
 2480  
 2490  
 2500  
 2510  
 2520  
 2530  
 2540  
 2550  
 2560  
 2570  
 2580  
 2590  
 2600  
 2610  
 2620  
 2630  
 2640  
 2650  
 2660  
 2670  
 2680  
 2690  
 2700  
 2710  
 2720  
 2730  
 2740  
 2750  
 2760  
 2770  
 2780  
 2790  
 2800  
 2810  
 2820  
 2830  
 2840  
 2850  
 2860  
 2870  
 2880  
 2890  
 2900  
 2910  
 2920  
 2930  
 2940  
 2950  
 2960  
 2970  
 2980  
 2990  
 3000  
 3010  
 3020  
 3030  
 3040  
 3050  
 3060  
 3070  
 3080  
 3090  
 3100  
 3110  
 3120  
 3130  
 3140  
 3150  
 3160  
 3170  
 3180  
 3190  
 3200  
 3210  
 3220  
 3230  
 3240  
 3250  
 3260  
 3270  
 3280  
 3290  
 3300  
 3310  
 3320  
 3330  
 3340  
 3350  
 3360  
 3370  
 3380  
 3390  
 3400  
 3410  
 3420  
 3430  
 3440  
 3450  
 3460  
 3470  
 3480  
 3490  
 3500  
 3510  
 3520  
 3530  
 3540  
 3550  
 3560  
 3570  
 3580  
 3590  
 3600  
 3610  
 3620  
 3630  
 3640  
 3650  
 3660  
 3670  
 3680  
 3690  
 3700  
 3710  
 3720  
 3730  
 3740  
 3750  
 3760  
 3770  
 3780  
 3790  
 3800  
 3810  
 3820  
 3830  
 3840  
 3850  
 3860  
 3870  
 3880  
 3890  
 3900  
 3910  
 3920  
 3930  
 3940  
 3950  
 3960  
 3970  
 3980  
 3990  
 4000  
 4010  
 4020  
 4030  
 4040  
 4050  
 4060  
 4070  
 4080  
 4090  
 4100  
 4110  
 4120  
 4130  
 4140  
 4150  
 4160  
 4170  
 4180  
 4190  
 4200  
 4210  
 4220  
 4230  
 4240  
 4250  
 4260  
 4270  
 4280  
 4290  
 4300  
 4310  
 4320  
 4330  
 4340  
 4350  
 4360  
 4370  
 4380  
 4390  
 4400  
 4410  
 4420  
 4430  
 4440  
 4450  
 4460  
 4470  
 4480  
 4490  
 4500  
 4510  
 4520  
 4530  
 4540  
 4550  
 4560  
 4570  
 4580  
 4590  
 4600  
 4610  
 4620  
 4630  
 4640  
 4650  
 4660  
 4670  
 4680  
 4690  
 4700  
 4710  
 4720  
 4730  
 4740  
 4750  
 4760  
 4770  
 4780

X  
 ACGACAT  
 |||||  
 ACGACAT  
 220 X

19. US-09-697-123B-24 (1-208)  
US-09-697-123B-4 Sequence 4, Application US/09697123B

Initial Score = 138 Optimized Score = 167 Significance = -0.21  
Residue Identity = 82% Matches = 172 Mismatches = 35  
Gaps = 1 Conservative Substitutions = 0

```
X      10      20      30      40      50      60      70
TCAAGCAGAAACCGCTACGACTGCGCCGGTGGCCCTTACAAAGTTAAACAAGAACTCGGGCTGAACACCG
|||||
TCAAGAGAAACCGCTACGACTGCGCCGGTGGCCCTTACAAAGTTAAACAAGAACTCGGGCTGATCTGCG
X      10      20      30      40      50      60      70
```

```
80      90      100      110      120      130      140
AGATTCGCCACCAACCAACCTCTGACCGAAGAGAGCTGCTGCCACCATCGAATATCTGCTGCGCTTTC
|||||
GCGATTCGATTCACGAGCTCGACCGCTGACCGAAGAGAGCTGCTGCCACCATCGAATATCTGCTGCGCTTTC
80      90      100      110      120      130      140
```

```
150      160      170      180      190      200
ACGAGGGGCGACCCCGCATGAGTATCCCGGCTGGCGCTGAGAGTCCGCGTGAAGACCGACAT
|||||
ACGAGGGGCGACCGATG-ACGTTCCGGGCGGACCGAGTTCGCTGAGAGACCGACGACAT
150      160      170      180      190      200
```

## 20. US-09-697-123B-24 (1-208)

US-09-697-123B-9 Sequence 9, Application US/09697123B

Initial Score = 136 Optimized Score = 160 Significance = -0.25  
Residue Identity = 76% Matches = 170 Mismatches = 38  
Gaps = 15 Conservative Substitutions = 0

```
X      10      20      30      40      50      60      70
TCAAGAGAAACCGCTACGACTGCGCCGGTGGCCCTTACAAAGTTAAACAAGAACTCGGGCTGAACACCG
|||||
TCAAGAGAAACCGCTACGACTGCGCCGGTGGCCCTTACAAAGTTAAACAAGAACTCGGGCTGACACGCGG
X      10      20      30      40      50      60      70
```

```
80      90      100      110      120      130      140
AGATTCGCCACCAACCAACCTCTGACCGAAGAGAGCTGCTGCCACCATCGAATATCTGCTGCGCTTTC
|||||
GCGAGCGGATTCAGCTCTGCTCCACCGTCCGAGAGAGAGCTGCTGCCACCATCGAATATCTGCTGCGCTTC
80      90      100      110      120      130      140
```

```
150      160      170      180      190      200
ACGAGG-----GGCAGCGCCACGATGAAGTCCCGGCTGGCTGAGAGTCCGGTGAAGACCG
|||||
ACCGCGCGCTACGATGCGCAGCGCCCGCTCATGTGACTGTCCCGGCGGATCGAGAGTGGCGTGAAGACCG
150      160      170      180      190      200      210
```

```
X
ACGACAT
|||||
ACGACAT
220
```

## 21. US-09-697-123B-24 (1-208)

US-09-697-123B-7 Sequence 7, Application US/09697123B

Initial Score = 123 Optimized Score = 161 Significance = -0.53  
Residue Identity = 78% Matches = 168 Mismatches = 40  
Gaps = 6 Conservative Substitutions = 0

```
X      10      20      30      40      50      60      70
TCAAGAGAAACCGCTACGACTGCGCCGGTGGCCCTTACAAAGTTAAACAAGAACTCGGGCTGAACACCG
|||||
TCAAGAGAAACCGCTACGACTGCGCCGGTGGCCCTTACAAAGTTAAACAAGAACTCGGGCTGATCTGCGG
X      10      20      30      40      50      60      70
```

```
80      90      100      110      120      130
AGATTCGCC-----AACCAACCAACCTCTGACCGAAGAGAGCTGCTGCCACCATCGAATATCTGCTG
|||||
CCAACCGGCTCTGCTGACTGCTCCACCGCTCAACGAGAAACGCTGCTGCGCACCATCGGCTGCTGCTG
80      90      100      110      120      130      140
```

```
140      150      160      170      180      190      200
GCTTGCACAGAGGGGACCGCCACGATGAAGTCCCGGCTGCGAGTGGCGGTGAGAGACCGACGACAT
|||||
GCTTGCACAGAGGGGACCGCCACGATGAAGTCCCGGCTGCGAGTGGCGGTGAGAGTGAAGTGAAGTGA
150      160      170      180      190      200      210
```

## 22. US-09-697-123B-24 (1-208)

US-09-697-123B-1 Sequence 18, Application US/09697123B

Initial Score = 119 Optimized Score = 161 Significance = -0.61  
Residue Identity = 79% Matches = 167 Mismatches = 41  
Gaps = 3 Conservative Substitutions = 0

```
X      10      20      30      40      50      60
TCAAGAGAAACCGCTACGACTGCGCCGGTGGCCCTTACAAAGTTAAACAAGAACTCGGGCTGA---ACA
|||||
TCAAGAGAAACCGCTACGACTGCGCCGGTGGCCCTTACAAAGTTAAACAAGAACTCGGGCTGATCTGCGG
X      10      20      30      40      50      60      70
```

```
70      80      90      100      110      120      130      140
CCGAGATTCGCCACCAACCAACCTCTGACCGAAGAGAGCTGCTGCCACCATCGAATATCTGCTGCGCT
|||||
ACCGGCGCGACACGACCTTCGACCGACCGCTGACCGAAGAGAGCTGCTGCCACCATCGAATATCTGCTGCGG
80      90      100      110      120      130      140
```

```
150      160      170      180      190      200
TTCACGAGGGGCGACCGCATGATGAAGTCCCGGCTGGCTGAGAGTCCGCGTGAAGACCGACGACAT
|||||
TTCACGAGGGGCGACCGCATGATGAAGTCCCGGCTGGCTGAGAGTCCGCGTGAAGTCCGCGTGAAGACG
150      160      170      180      190      200      210
```

## 23. US-09-697-123B-24 (1-208)

US-09-697-123B-2 Sequence 21, Application US/09697123B

Initial Score = 111 Optimized Score = 166 Significance = -0.78  
Residue Identity = 80% Matches = 173 Mismatches = 35  
Gaps = 6 Conservative Substitutions = 0

```
X      10      20      30      40      50      60      70
TCAAGAGAAACCGCTACGACTGCGCCGGTGGCCCTTACAAAGTTAAACAAGAACTCGGGCTGAACACCG
|||||
TCAAGAGAAACCGCTACGACTGCGCCGGTGGCTGAGTGAAGTGAAGAAAGTGGCGTGGCGGGA
X      10      20      30      40      50      60      70
```

```
80      90      100      110      120      130
AGATTCGCC-----AACCAACCAACCTCTGACCGAAGAGAGCTGCTGCCACCATCGAATATCTGCTG
|||||
CCAATTCGGCTTACGATGACACCAACCAACCTCTGACCGAAGAGAGCTGCTGCCACCATCGAATATCTGCTG
80      90      100      110      120      130      140
```

```
140      150      160      170      180      190      200
GCTTGCACAGAGGGGACCGCCACGATGAAGTCCCGGCTGCGAGTGGCGGTGAGAGACCGACGACAT
|||||
GCTTGCACAGAGGGGACCGCCACGATGAAGTCCCGGCTGCGAGTGGCGGTGAGAGTGAAGTGAAGTGA
150      160      170      180      190      200      210
```

## 24. US-09-697-123B-24 (1-208)

US-09-697-123B-1 Sequence 14, Application US/09697123B

Initial Score = 109 Optimized Score = 171 Significance = -0.82  
Residue Identity = 83% Matches = 178 Mismatches = 30  
Gaps = 6 Conservative Substitutions = 0

```
X      10      20      30      40      50      60      70
TCAAGAGAAACCGCTACGACTGCGCCGGTGGCCCTTACAAAGTTAAACAAGAACTCGGGCTGAACACCG
|||||
TCAAGAGAAACCGCTACGACTGCGCCGGTGGCCCTTACAAAGTTAAACAAGAACTCGGGCTGCGGGGCGG
X      10      20      30      40      50      60      70
```

```
      80      90      100      110      120      130
AGAAATCG-----CCAAACACACGACCCCTGACCGAAGAGGAGCGTCCGCGCACCATCGAATACCTGTGTC
      |      |      |      |      |      |      |      |      |      |      |
CCGACTCGGCGCTACCCGCTCTGACACGCTGACGAGCGGATGTCTGCGCCACCATCGAGTACCTGTGTC
      80      90      100      110      120      130      140
140      150      160      170      180      190      200
GCTTGCACGACGCGCACGCGCATGTAAGATCCCGGTGCGCTGAGAGTGCCTGCGAGACCGAGACAT
      |      |      |      |      |      |      |      |      |      |      |
GCCTGCACGACGCGCGCATGATGACGCTTCCCGCGCGCTGAGAGTGCCGGTGGAGACCGAGACAT
      150      160      170      180      190      200      210
X
TCAGGAGAGAAGCGCTACGA
|||||
TCAGGAGAGAAGCGCTACGA
X
10
X
```

25. US-09-697-123B-24 (1-208)  
US-09-697-123B-2 Sequence 25. Application US/09697123B

Initial Score	19	Optimized Score	19	Significance
Residue Identity	100%	Matches	19	Mismatches
Gaps	0	Conservative Substitutions	-	0

X  
TCAGGAGAGAAGCGCTACGA  
|||||  
TCAGGAGAGAAGCGCTACGA  
X  
10  
X

> 0 <  
0110 IntelliGenetics  
> 0 <  
FastDB - Fast Pairwise Comparison of Sequences  
Release 5.4

Results file us-09-697-123b-23.res made by shanley on Wed 13 Nov 102 14:13:07 PST.

Query sequence being compared: US-09-697-123B-23 (1-208)  
Number of sequences searched: 26  
Number of scores above cutoff: 25

Results of the initial comparison of US-09-697-123B-23 (1-208) with:  
File: US09697123B.seq

100-  
N -  
U 50-  
M -  
B -  
E -  
R -  
O -  
F 10-  
S -  
E 5-  
U -  
N -  
C -  
S -  
E 0-  
SCORE 0 23 46 69 92 116 139 162 185 208  
STDDEV -2 -1 0 1

PARAMETERS  
Similarity matrix Unitary 1 K-tuple 4  
Mismatch penalty 5.00 Joining penalty 30  
Gap penalty 0.33 Window size 205  
Gap size penalty 12  
Cutoff score 0  
Randomization group 0

SEARCH STATISTICS

Scores: Mean 152 Median 177 Standard Deviation 48.36  
Times: CPU 00:00:00.00 Total Elapsed 00:00:00.00

Number of residues: 5077  
Number of sequences searched: 26  
Number of scores above cutoff: 25

The scores below are sorted by initial score.  
Significance is calculated based on initial score.

A 100% identical sequence to the query sequence was found:

Sequence Name	Description	Init. Opt.	Length	Score	Sig.	Frame
---------------	-------------	------------	--------	-------	------	-------

1. US-09-697-123B-2 Sequence 23, Application 208 208 208 1.16 0  
The list of other best scores is:

Sequence Name	Description	Length	Score	Init. Opt.	Sig.	Frame
2. US-09-697-123B-1 Sequence 15, Application 208		188	188	0.74	0	
3. US-09-697-123B-2 Sequence 2, Application U 208		185	185	0.68	0	
4. US-09-697-123B-3 Sequence 3, Application U 208		183	183	0.64	0	
5. US-09-697-123B-1 Sequence 16, Application U 208		182	182	0.62	0	
6. US-09-697-123B-6 Sequence 6, Application U 208		182	182	0.62	0	
7. US-09-697-123B-5 Sequence 5, Application U 208		182	182	0.62	0	
8. US-09-697-123B-2 Sequence 22, Application 208		181	181	0.60	0	
9. US-09-697-123B-1 Sequence 1, Application U 208		181	181	0.60	0	
10. US-09-697-123B-1 Sequence 11, Application 208		180	180	0.58	0	
11. US-09-697-123B-1 Sequence 10, Application 208		179	179	0.56	0	
12. US-09-697-123B-8 Sequence 8, Application U 208		177	177	0.52	0	
13. US-09-697-123B-1 Sequence 19, Application 208		176	176	0.50	0	
14. US-09-697-123B-1 Sequence 17, Application 208		176	176	0.50	0	
15. US-09-697-123B-2 Sequence 14, Application 208		173	173	0.43	0	
16. US-09-697-123B-1 Sequence 24, Application 207		172	172	0.00	0	
17. US-09-697-123B-4 Sequence 4, Application U 207		148	175	-0.08	0	
18. US-09-697-123B-2 Sequence 20, Application 205		147	171	-0.10	0	
19. US-09-697-123B-9 Sequence 9, Application U 223		138	163	-0.29	0	
20. US-09-697-123B-1 Sequence 13, Application 223		136	159	-0.33	0	
21. US-09-697-123B-1 Sequence 14, Application 214		132	169	-0.41	0	
22. US-09-697-123B-2 Sequence 21, Application 214		117	161	-0.72	0	
23. US-09-697-123B-7 Sequence 7, Application U 214		117	155	-0.72	0	
24. US-09-697-123B-1 Sequence 18, Application 211		113	165	-0.81	0	
25. US-09-697-123B-2 Sequence 25, Application 19		19	19	-2.75	0	

1. US-09-697-123B-23 (1-208)  
US-09-697-123B-2 Sequence 23, Application US/09697123B

Initial Score - 208 Optimized Score - 208 Significance - 1.16  
Residue Identity - 100% Matches - 208 Mismatches - 0  
Gaps - 0 Conservative Substitutions - 0

X 10 20 30 40 50 60 70  
TCAAGAGAAAGCGCTACGACCTGGCGGGTGTCTGCTCAAGTCAAGAAAGCTGGCTTGCACGGCG  
|||||  
TCAAGAGAAAGCGCTACGACCTGGCGGGTGTCTGCTCAAGTCAAGAAAGCTGGCTTGCACGGCG  
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140  
GTGAGCGGATCAGCAGCTCGACGCTGACCGAAGAGAGAGCGTCGTCACCATGAGTACTGTCGGCGTGC  
|||||  
GTGAGCGGATCAGCAGCTCGACGCTGACCGAAGAGAGAGCGTCGTCACCATGAGTACTGTCGGCGTGC  
X 80 90 100 110 120 130 140

150 160 170 180 190 200  
ATGAGGTCAGTACGATGACGATGACCGCTTCAGGTGCGCGAGCGGACGATGAGTACGATGACGAT  
|||||  
ATGAGGTCAGTACGATGACGATGACCGCTTCAGGTGCGCGAGCGGATGAGTACGATGACGAT  
X 150 160 170 180 190 200

2. US-09-697-123B-23 (1-208)  
US-09-697-123B-1 Sequence 15, Application US/09697123B

Initial Score - 188 Optimized Score - 188 Significance - 0.74  
Residue Identity - 90% Matches - 188 Mismatches - 20  
Gaps - 0 Conservative Substitutions - 0

X 10 20 30 40 50 60 70  
TCAAGAGAAAGCGCTACGACCTGGCGGGTGTCTGCTTCAAGTCAAGAAAGCTGGCTTGCACGGCG  
|||||  
TCAAGAGAAAGCGCTACGACCTGGCGGGTGTCTGCTTCAAGTCAAGAAAGCTGGCTTGCACGGCG  
X 10 20 30 40 50 60 70





Sequence Identity =	77%	Matches =	173	Mismatches =	35
	15	Conservative Substitutions			0

  

X	10	20	30	40	50	60	70
CAAGAAGAGCGCTACGACCTCGGCCCCGGGTTGGTCTTACACAGTCAACACGAAGCTCGGGGTTGCACGCGG							







> 0 <  
0110 Intelligenetics  
> 0 <  
FastDB - Fast Pairwise Comparison of Sequences  
Release 5.4

Results file us-09-697-123b-22.res made by shanley on Wed 13 Nov 102 14:12:29-PST.

Query sequence being compared: US-09-697-123B-22 (1-208)  
Number of sequences searched: 26  
Number of scores above cutoff: 25

Results of the initial comparison of US-09-697-123B-22 (1-208) with:  
File: US09697123B.seq

100-  
N -  
U -  
M -  
B -  
E -  
R -  
O -  
F -  
S -  
E -  
U -  
N -  
C -  
S -  
SCORE 0 231 46 69 92 116 139 162 185 208  
STDEV -2 -1 0 1

PARAMETERS  
Similarity matrix Unitary K-tuple  
Mismatch penalty 1 Joining penalty 4  
Gap penalty 5.00 Window size 30  
Gap size penalty 0.33 205  
Cutoff score 12  
Randomization group 0

SEARCH STATISTICS  
Scores: Mean 155 Median 178 Standard Deviation 51.71  
Times: CPU 00:00:00.00 Total Elapsed 00:00:00.00  
Number of residues: 5077  
Number of sequences searched: 26  
Number of scores above cutoff: 25

The scores below are sorted by initial score.  
Significance is calculated based on initial score.

A 100% identical sequence to the query sequence was found:

Sequence Name Description Init. Opt. Length Score Score Sig. Frame

1. US-09-697-123B-2 Sequence 22, Application 208 208 208 1.02 0  
The list of other best scores is:

Sequence Name Description Length Score Score Init. Opt. Sig. Frame  
-----  
2. US-09-697-123B-1 Sequence 16, Application 208 207 207 1.01 0  
3. US-09-697-123B-6 Sequence 6, Application 208 207 207 1.01 0  
4. US-09-697-123B-5 Sequence 5, Application 208 207 207 1.01 0  
5. US-09-697-123B-1 Sequence 10, Application 208 185 185 0.58 0  
6. US-09-697-123B-1 Sequence 11, Application 208 184 184 0.56 0  
7. US-09-697-123B-3 Sequence 3, Application 208 184 184 0.56 0  
8. US-09-697-123B-1 Sequence 15, Application 208 183 183 0.54 0  
9. US-09-697-123B-2 Sequence 23, Application 208 181 181 0.50 0  
10. US-09-697-123B-1 Sequence 19, Application 208 181 181 0.50 0  
11. US-09-697-123B-8 Sequence 8, Application 208 181 181 0.50 0  
12. US-09-697-123B-1 Sequence 1, Application 208 181 181 0.50 0  
13. US-09-697-123B-2 Sequence 2, Application 208 180 180 0.48 0  
14. US-09-697-123B-1 Sequence 17, Application 208 177 177 0.43 0  
15. US-09-697-123B-2 Sequence 24, Application 208 172 172 0.33 0  
16. US-09-697-123B-2 Sequence 20, Application 205 160 170 0.10 0  
17. US-09-697-123B-1 Sequence 12, Application 207 156 175 0.02 0  
18. US-09-697-123B-4 Sequence 4, Application 207 153 178 -0.04 0  
19. US-09-697-123B-9 Sequence 9, Application 223 143 159 -0.23 0  
20. US-09-697-123B-1 Sequence 13, Application 223 136 161 -0.37 0  
21. US-09-697-123B-7 Sequence 7, Application 214 120 158 -0.37 0  
22. US-09-697-123B-1 Sequence 18, Application 211 114 158 -0.68 0  
23. US-09-697-123B-1 Sequence 14, Application 214 109 167 -0.79 0  
24. US-09-697-123B-2 Sequence 21, Application 214 108 161 -0.89 0  
25. US-09-697-123B-2 Sequence 25, Application 19 19 19 -2.63 0  
\*\*\*\* 2 standard deviations below mean \*\*\*\*

1. US-09-697-123B-22 (1-208)  
US-09-697-123B-2 Sequence 22, Application US/09697123B  
Initial Score - 208 Optimized Score - 208 Significance - 1.02  
Residue Identity - 100% Matches - 208 Mismatches - 0  
Gaps - 0 Conservative Substitutions - 0

X 10 20 30 40 50 60 70  
TCAAGAGAAAGCGCTACGACCTGCGCCGCGTGGCTATTAAGTCAACAAAGAGCGGCGATGTCG  
TCAAGAGAAAGCGCTACGACCTGCGCCGCGTGGCTATTAAGTCAACAAAGAGCGGCGATGTCG  
TCAAGAGAAAGCGCTACGACCTGCGCCGCGTGGCTATTAAGTCAACAAAGAGCGGCGATGTCG  
X 10 20 30 40 50 60 70

CGGAGCCCATCTACGTCGTGAGCGCTGACCGAAGAGAGAGCGTGGCCACCATCGAATATCTGTCGCTGC  
CGGAGCCCATCTACGTCGTGAGCGCTGACCGAAGAGAGAGCGTGGCCACCATCGAATATCTGTCGCTGC  
CGGAGCCCATCTACGTCGTGAGCGCTGACCGAAGAGAGAGCGTGGCCACCATCGAATATCTGTCGCTGC  
X 80 90 100 110 120 130 140

ACGAGGCTCAGACCGATGATGCTTCGCGCGCGCTGAGAGCTCCGGTGAACACCGACACAT X  
ACGAGGCTCAGACCGATGATGCTTCGCGCGCGCTGAGAGCTCCGGTGAACACCGACACAT X  
ACGAGGCTCAGACCGATGATGCTTCGCGCGCGCTGAGAGCTCCGGTGAACACCGACACAT X  
X 150 160 170 180 190 200

2. US-09-697-123B-22 (1-208)  
US-09-697-123B-1 Sequence 16, Application US/09697123B

Initial Score - 207 Optimized Score - 207 Significance - 1.01  
Residue Identity - 99% Matches - 207 Mismatches - 1  
Gaps - 0 Conservative Substitutions - 0

X 10 20 30 40 50 60 70  
TCAAGAGAAAGCGCTACGACCTGCGCCGCGTGGCTATTAAGTCAACAAAGAGCGGCGATGTCG  
TCAAGAGAAAGCGCTACGACCTGCGCCGCGTGGCTATTAAGTCAACAAAGAGCGGCGATGTCG  
TCAAGAGAAAGCGCTACGACCTGCGCCGCGTGGCTATTAAGTCAACAAAGAGCGGCGATGTCG  
X 10 20 30 40 50 60 70

8. US-09-697-123B-22 (1-208)



## 14. US-09-697-123B-1 Sequence 17, Application US/09697123B

Initial Score = 177 Optimized Score = 177 Significance = 0.43  
Residue Identity = 85% Matches = 177 Mismatches = 31  
Gaps = 0 Conservative Substitutions = 0

```
150 160 170 180 190 200
ACGAGGCTCAGACCATGATGCTGTCGGGCGGCTCGAGGTGCGGTGAACCGACGACAT
|||||
ACGAGGCTCAGACCATGATGCTGTCGGGCGGCTCGAGGTGCGGTGAACCGACGACAT
150 160 170 180 190 200
X
```

```
150 160 170 180 190 200
TCAAGAGAACCGCTACGACCTGCGCGGCTGCGCTATTAAGTCAACAGAGCTCGGCTGATGTCG
|||||
TCAAGAGAACCGCTACGACCTGCGCGGCTGCGCTATTAAGTCAACAGAGCTCGGCTGATGTCG
150 160 170 180 190 200
X
```

```
150 160 170 180 190 200
GCGAGCCATACGCTGCTGACGCTGACCGAAGAGAGCTGCGCCACCATGATATGCTGCTGTC
|||||
GCGAGCCATACGCTGCTGACGCTGACCGAAGAGAGCTGCGCCACCATGATATGCTGCTGTC
150 160 170 180 190 200
X
```

```
150 160 170 180 190 200
ACGAGGCTCAGACCATGATGCTGTCGGGCGGCTCGAGGTGCGGTGAACCGACGACAT
|||||
ACGAGGCTCAGACCATGATGCTGTCGGGCGGCTCGAGGTGCGGTGAACCGACGACAT
150 160 170 180 190 200
X
```

## 15. US-09-697-123B-22 (1-208) Sequence 24, Application US/09697123B

Initial Score = 172 Optimized Score = 172 Significance = 0.33  
Residue Identity = 82% Matches = 172 Mismatches = 36  
Gaps = 0 Conservative Substitutions = 0

```
150 160 170 180 190 200
TCAAGAGAACCGCTACGACCTGCGCGGCTGCGCTATTAAGTCAACAGAGCTCGGCTGATGTCG
|||||
TCAAGAGAACCGCTACGACCTGCGCGGCTGCGCTATTAAGTCAACAGAGCTCGGCTGATGTCG
150 160 170 180 190 200
X
```

```
150 160 170 180 190 200
GCGAGCCATACGCTGCTGACGCTGACCGAAGAGAGCTGCGCCACCATGATATGCTGCTGTC
|||||
GCGAGCCATACGCTGCTGACGCTGACCGAAGAGAGCTGCGCCACCATGATATGCTGCTGTC
150 160 170 180 190 200
X
```

```
150 160 170 180 190 200
ACGAGGCTCAGACCATGATGCTGTCGGGCGGCTCGAGGTGCGGTGAACCGACGACAT
|||||
ACGAGGCTCAGACCATGATGCTGTCGGGCGGCTCGAGGTGCGGTGAACCGACGACAT
150 160 170 180 190 200
X
```

## 16. US-09-697-123B-22 (1-208) Sequence 20, Application US/09697123B

Initial Score = 160 Optimized Score = 176 Significance = 0.10  
Residue Identity = 84% Matches = 176 Mismatches = 29  
Gaps = 3 Conservative Substitutions = 0

```
150 160 170 180 190 200
TCAAGAGAACCGCTACGACCTGCGCGGCTGCGCTATTAAGTCAACAGAGCTCGGCTGATGTCG
|||||
TCAAGAGAACCGCTACGACCTGCGCGGCTGCGCTATTAAGTCAACAGAGCTCGGCTGATGTCG
150 160 170 180 190 200
X
```

```
150 160 170 180 190 200
GCGAGCCATACGCTGCTGACGCTGACCGAAGAGAGCTGCGCCACCATGATATGCTGCTGTC
|||||
GCGAGCCATACGCTGCTGACGCTGACCGAAGAGAGCTGCGCCACCATGATATGCTGCTGTC
150 160 170 180 190 200
X
```

```
150 160 170 180 190 200
ACGAGGCTCAGACCATGATGCTGTCGGGCGGCTCGAGGTGCGGTGAACCGACGACAT
|||||
ACGAGGCTCAGACCATGATGCTGTCGGGCGGCTCGAGGTGCGGTGAACCGACGACAT
150 160 170 180 190 200
X
```

## 17. US-09-697-123B-22 (1-208) Sequence 12, Application US/09697123B

Initial Score = 156 Optimized Score = 175 Significance = 0.02  
Residue Identity = 86% Matches = 180 Mismatches = 27  
Gaps = 1 Conservative Substitutions = 0

```
150 160 170 180 190 200
TCAAGAGAACCGCTACGACCTGCGCGGCTGCGCTATTAAGTCAACAGAGCTCGGCTGATGTCG
|||||
TCAAGAGAACCGCTACGACCTGCGCGGCTGCGCTATTAAGTCAACAGAGCTCGGCTGATGTCG
150 160 170 180 190 200
X
```

```
150 160 170 180 190 200
GCGAGCCATACGCTGCTGACGCTGACCGAAGAGAGCTGCGCCACCATGATATGCTGCTGTC
|||||
GCGAGCCATACGCTGCTGACGCTGACCGAAGAGAGCTGCGCCACCATGATATGCTGCTGTC
150 160 170 180 190 200
X
```

```
150 160 170 180 190 200
ACGAGGCTCAGACCATGATGCTGTCGGGCGGCTCGAGGTGCGGTGAACCGACGACAT
|||||
ACGAGGCTCAGACCATGATGCTGTCGGGCGGCTCGAGGTGCGGTGAACCGACGACAT
150 160 170 180 190 200
X
```

## 18. US-09-697-123B-22 (1-208) Sequence 4, Application US/09697123B

Initial Score = 153 Optimized Score = 178 Significance = -0.04  
Residue Identity = 87% Matches = 183 Mismatches = 24  
Gaps = 1 Conservative Substitutions = 0

```
150 160 170 180 190 200
TCAAGAGAACCGCTACGACCTGCGCGGCTGCGCTATTAAGTCAACAGAGCTCGGCTGATGTCG
|||||
TCAAGAGAACCGCTACGACCTGCGCGGCTGCGCTATTAAGTCAACAGAGCTCGGCTGATGTCG
150 160 170 180 190 200
X
```

```
150 160 170 180 190 200
GCGAGCCATACGCTGCTGACGCTGACCGAAGAGAGCTGCGCCACCATGATATGCTGCTGTC
|||||
GCGAGCCATACGCTGCTGACGCTGACCGAAGAGAGCTGCGCCACCATGATATGCTGCTGTC
150 160 170 180 190 200
X
```

```
150 160 170 180 190 200
ACGAGGCTCAGACCATGATGCTGTCGGGCGGCTCGAGGTGCGGTGAACCGACGACAT
|||||
ACGAGGCTCAGACCATGATGCTGTCGGGCGGCTCGAGGTGCGGTGAACCGACGACAT
150 160 170 180 190 200
X
```

## 19. US-09-697-123B-22 (1-208) Sequence 9, Application US/09697123B

Initial Score = 143 Optimized Score = 169 Significance = -0.23  
Residue Identity = 80% Matches = 179 Mismatches = 29  
Gaps = 15 Conservative Substitutions = 0

X  
10 20 30 40 50 60 70  
TCAAGAGAGAGCGCTACGACCGCCCGCGTGGTGGTATAGGTCAAGAAAGCTCGCATGTG  
|||||  
TCAAGAGAGAGCGCTACGACCGCCCGCGTGGTGGTATAGGTCAAGAAAGCTCGCATGTG  
X 10 20 30 40 50 60 70  
CGGAGCCCATCAGCTGCTCAAGCGTGAAGAGAGAGAGTGTGCGGATGATGCTGTGCGGCTTC  
80 90 100 110 120 130 140  
CGGAGCCCATCAGCTGCTCAAGCGTGAAGAGAGAGAGTGTGCGGATGATGCTGTGCGGCTTC  
80 90 100 110 120 130 140  
ACGAGGCTCAGA-----CCAGCATGATGCTTCCGGCGCGCTGAGGAGCGGAGAAACCG  
|||  
ACGAGGCTCAGA-----CCAGCATGATGCTTCCGGCGCGCTGAGGAGCGGAGAAACCG  
150 160 170 180 190 200 210  
X  
ACGACAT  
|||||  
ACGACAT  
220 X

20. US-09-697-123B-22 (1-208)  
US-09-697-123B-1 Sequence 13, Application US/09697123B

Initial Score = 136 Optimized Score = 161 Significance = -0.37  
Residue Identity = 78% Matches = 175 Mismatches = 33  
Gaps = 15 Conservative Substitutions = 0  
X  
10 20 30 40 50 60 70  
TCAAGAGAGAGCGCTACGACCGCCCGCGTGGTGGTATAGGTCAAGAAAGCTCGCATGTG  
|||||  
TCAAGAGAGAGCGCTACGACCGCCCGCGTGGTGGTATAGGTCAAGAAAGCTCGCATGTG  
X 10 20 30 40 50 60 70  
CGGAGCCCATCAGCTGCTCAAGCGTGAAGAGAGAGTGTGCGGATGATGCTGTGCGGCTTC  
80 90 100 110 120 130 140  
CGGAGCCCATCAGCTGCTCAAGCGTGAAGAGAGAGTGTGCGGATGATGCTGTGCGGCTTC  
80 90 100 110 120 130 140  
ACGAGGCTCAGA-----CCAGCATGATGCTTCCGGCGCGCTGAGGAGCGGAGAAACCG  
|||  
ACGAGGCTCAGA-----CCAGCATGATGCTTCCGGCGCGCTGAGGAGCGGAGAAACCG  
150 160 170 180 190 200 210  
X  
ACGACAT  
|||||  
ACGACAT  
220 X

21. US-09-697-123B-22 (1-208)  
US-09-697-123B-7 Sequence 7, Application US/09697123B

Initial Score = 120 Optimized Score = 158 Significance = -0.68  
Residue Identity = 77% Matches = 165 Mismatches = 43  
Gaps = 6 Conservative Substitutions = 0  
X  
10 20 30 40 50 60 70  
TCAAGAGAGAGCGCTACGACCGCCCGCGTGGTGGTATAGGTCAAGAAAGCTCGCATGTG  
|||||  
TCAAGAGAGAGCGCTACGACCGCCCGCGTGGTGGTATAGGTCAAGAAAGCTCGCATGTG  
X 10 20 30 40 50 60 70  
CGGAGCCCATCAGCTGCTCAAGCGTGAAGAGAGAGTGTGCGGATGATGCTGTGCGGCTTC  
80 90 100 110 120 130  
CGGAGCCCATCAGCTGCTCAAGCGTGAAGAGAGAGTGTGCGGATGATGCTGTGCGGCTTC  
|||

CCAAACCGGCTGTGTGACTGTCCACACGCTTACGAGAGAAAGCTGTGCGCACCATCGGGTACCTGTGC  
80 90 100 110 120 130 140  
GCTTGCAGAGAGGCGTCAAGACCATGATGCTTCCGGCGCGTGAAGTGTGCGTGAAGACGACAT  
|||  
GCTTGCAGAGAGGCGTCAAGACCATGATGCTTCCGGCGCGTGAAGTGTGCGTGAAGACGACAT  
150 160 170 180 190 200 210 X

22. US-09-697-123B-22 (1-208)  
US-09-697-123B-1 Sequence 18, Application US/09697123B

Initial Score = 114 Optimized Score = 158 Significance = -0.79  
Residue Identity = 77% Matches = 164 Mismatches = 44  
Gaps = 3 Conservative Substitutions = 0  
X  
10 20 30 40 50 60 70  
TCAAGAGAGAGCGCTACGACCGCCCGCGTGGTGGTATAGGTCAAGAAAGCTCGCATGTG  
|||||  
TCAAGAGAGAGCGCTACGACCGCCCGCGTGGTGGTATAGGTCAAGAAAGCTCGCATGTG  
X 10 20 30 40 50 60 70  
CGGAGCCCATCAGCTGCTCAAGCGTGAAGAGAGAGTGTGCGGATGATGCTGTGCGGCTTC  
80 90 100 110 120 130 140  
CGGAGCCCATCAGCTGCTCAAGCGTGAAGAGAGAGTGTGCGGATGATGCTGTGCGGCTTC  
80 90 100 110 120 130 140  
ACGAGGCTCAGA-----CCAGCATGATGCTTCCGGCGCGCTGAGGAGCGGAGAAACCG  
|||  
ACGAGGCTCAGA-----CCAGCATGATGCTTCCGGCGCGCTGAGGAGCGGAGAAACCG  
150 160 170 180 190 200 210  
X  
ACGACAT  
|||||  
ACGACAT  
220 X

23. US-09-697-123B-22 (1-208)  
US-09-697-123B-1 Sequence 14, Application US/09697123B

Initial Score = 109 Optimized Score = 167 Significance = -0.89  
Residue Identity = 81% Matches = 174 Mismatches = 34  
Gaps = 6 Conservative Substitutions = 0  
X  
10 20 30 40 50 60 70  
TCAAGAGAGAGCGCTACGACCGCCCGCGTGGTGGTATAGGTCAAGAAAGCTCGCATGTG  
|||||  
TCAAGAGAGAGCGCTACGACCGCCCGCGTGGTGGTATAGGTCAAGAAAGCTCGCATGTG  
X 10 20 30 40 50 60 70  
CGGAGCCCATCAGCTGCTCAAGCGTGAAGAGAGAGTGTGCGGATGATGCTGTGCGGCTTC  
80 90 100 110 120 130  
CGGAGCCCATCAGCTGCTCAAGCGTGAAGAGAGAGTGTGCGGATGATGCTGTGCGGCTTC  
80 90 100 110 120 130  
ACGAGGCTCAGA-----CCAGCATGATGCTTCCGGCGCGCTGAGGAGCGGAGAAACCG  
|||  
ACGAGGCTCAGA-----CCAGCATGATGCTTCCGGCGCGCTGAGGAGCGGAGAAACCG  
150 160 170 180 190 200 210  
X  
ACGACAT  
|||||  
ACGACAT  
220 X

24. US-09-697-123B-22 (1-208)  
US-09-697-123B-2 Sequence 21, Application US/09697123B

Initial Score = 108 Optimized Score = 161 Significance = -0.91  
Residue Identity = 78% Matches = 168 Mismatches = 40  
Gaps = 6 Conservative Substitutions = 0  
X  
10 20 30 40 50 60 70  
TCAAGAGAGAGCGCTACGACCGCCCGCGTGGTGGTATAGGTCAAGAAAGCTCGCATGTG  
|||||  
TCAAGAGAGAGCGCTACGACCGCCCGCGTGGTGGTATAGGTCAAGAAAGCTCGCATGTG  
X 10 20 30 40 50 60 70  
CGGAGCCCATCAGCTGCTCAAGCGTGAAGAGAGAGTGTGCGGATGATGCTGTGCGGCTTC  
80 90 100 110 120 130  
CGGAGCCCATCAGCTGCTCAAGCGTGAAGAGAGAGTGTGCGGATGATGCTGTGCGGCTTC  
80 90 100 110 120 130  
ACGAGGCTCAGA-----CCAGCATGATGCTTCCGGCGCGCTGAGGAGCGGAGAAACCG  
|||  
ACGAGGCTCAGA-----CCAGCATGATGCTTCCGGCGCGCTGAGGAGCGGAGAAACCG  
150 160 170 180 190 200 210 X





Results file us-09-697-123b-21.res made by shanley on wed 13 Nov 102 14:11:41-PST.

results of the initial comparison of US-09-697-123B-21 (1-214) with  
File : US09697123B.seq

SCORE	0	24	48	71	95	119	143	166	190	214
STDEV	-2		-1		0		1		2	

## PARAMETERS

	Unary	K-tuple
Similarity matrix	1	4
Mismatch penalty	1	30
Gap penalty	5.00	Window size
Gap size penalty	0.33	205
Cutoff score	12	
Randomization group	0	

## SEARCH STATISTICS

Scores:	Mean	Median	Standard Deviation
	112	112	42.66
Times:	CPU	Total Elapsed	
	00:00:00.00	00:00:00.00	

Number of residues:	5077
Number of sequences searched:	26
Number of scores above cutoff:	25

The scores below are sorted by initial score. Significance is calculated based on initial score.

A 100% identical sequence to the query sequence was found:

Sequence Name	Description	Init. Opt.	Length	Score	Score	Sig.	Frame
---------------	-------------	------------	--------	-------	-------	------	-------

1. US-09-697-123B-2 Sequence 21, Application	214	214	2.39	0
The list of other best scores is:				

Sequence Name	Description	Init. Opt.	Length	Score	Sig. Frame
---------------	-------------	------------	--------	-------	------------

	Standard deviation	Above mean	****
2. US-09-697-123B-7 Sequence 7, Application U	214	197	1.99
3. US-09-697-123B-1 Sequence 14, Application	214	171	1.71
**** 0 standard deviation from mean ****			1.38
4. US-09-697-123B-1 Sequence 10, Application	208	139	1.70
5. US-09-697-123B-1 Sequence 11, Application	208	138	1.69
6. US-09-697-123B-1 Sequence 15, Application	208	125	1.77
7. US-09-697-123B-8 Sequence 8, Application U	208	122	1.70
8. US-09-697-123B-2 Sequence 23, Application	208	117	1.61
9. US-09-697-123B-3 Sequence 15, Application	208	112	1.72
10. US-09-697-123B-1 Sequence 3, Application U	208	112	1.66
11. US-09-697-123B-2 Sequence 2, Application U	208	112	1.67
12. US-09-697-123B-2 Sequence 24, Application U	208	111	1.66
13. US-09-697-123B-1 Sequence 17, Application	208	111	1.69
14. US-09-697-123B-1 Sequence 18, Application	211	111	1.70
15. US-09-697-123B-2 Sequence 20, Application	205	109	1.59
16. US-09-697-123B-1 Sequence 16, Application	208	109	1.62
17. US-09-697-123B-1 Sequence 6, Application U	208	109	1.62
18. US-09-697-123B-5 Sequence 5, Application U	208	109	1.62
19. US-09-697-123B-2 Sequence 1, Application U	208	109	1.70
20. US-09-697-123B-2 Sequence 22, Application	208	108	1.61
21. US-09-697-123B-2 Sequence 12, Application	207	106	1.58
22. US-09-697-123B-1 Sequence 9, Application U	203	83	1.52
23. US-09-697-123B-4 Sequence 4, Application U	207	82	1.57
24. US-09-697-123B-1 Sequence 19, Application	223	81	1.55
**** 2 standard deviations below mean ****			-0.73
25. US-09-697-123B-2 Sequence 25, Application	19	19	-2.18

1. US-09-697-123B-21 (1-214)  
US-09-697-123B-2 Sequence 21, Application US/09697123B

Initial score	-	214	Optimized Score	-	214	Significance	-	2.39
Residue Identity	-	100%	Matches	-	214	Mismatches	-	0
Gaps	-	0	Conservative Substitutions	-			-	0

[illegible][illegible]

150	160	170	180	190	200	210	X
CTGCACGAGGGCCGAGACCAAGATGACCGCCCCGGCGGCGTGCAGATGCCGATGATGTGACACACT							
CTGCACGAGGGCCGAGACCAAGATGACCGCCCCGGCGGCGTGCAGATGCCGATGATGTGACACACT							
150	160	170	180	190	200	210	X

2. US-09-697-123B-21 (1-214)  
US-09-697-123B-7 Sequence 7, Application US/09697123B

Initial Score	-	197	Optimized Score	-	197	Significance	-	1.99
Residue Identity	=	92%	Matches	-	197	Mismatches	-	17
Gaps	=	0	Conservative Substitutions	-			-	0

X  
10 20 30 40 50 60 70

TCAAGSAGAAGCGCTACGATTTGGCCCGGTGGTCGTGCATCAAGGTGAACAAGCATGGCCCTGGCGGCCA  
TCAAAGSAGAAGCGCTACGACTTGCCCCGCTGGCGCGGTACAAGGTGAACAAGCATTGGCTTTGGCGCGGTG

8. US-09-697-123B-21 (1-214)

CCAATCCGGCCTAGGTGACACCAACCCTCACGAGAAACCTCTTCCCACCATCGAGTAACCTGGTC  
80           90          100         110         120         130         140  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
ACACCGGTCGCCGATCGACGACACACTGTGACCAAGSAGCTCGTGCCACCATGAGTAACCTGGTCC

70 80 90 100 110 120 130

150 160 170 180 190 200 210 X  
CCCTTCAGAGAGCCGACAGACGATGACCCCGCGCGGTGAGCTGGCGGTGATGTGACGACAT  
|||||  
GCGTGCAGAGAGCCGACAGACGATGACCCCGCGCGGTGAGCTGGCGGTGATGTGACGACAT  
140 150 160 170 180 190 200 X

## 14. US-09-697-123b-21 (1-214)

US-09-697-123b-1 Sequence 18, Application US/09697123b

Initial Score = 111 Optimized Score = 170 Significance = -0.02  
Residue Identity = 82% Matches = 176 Mismatches = 35  
Gaps = 3 Conservative Substitutions = 0

X 10 20 30 40 50 60 70  
TCAGAGAGAGCCGCTACGATCTGGCCGCGGTGCGTCAAGGTGACAGAGAGAGCTGGCGCGCA  
|||||  
TCAGAGAGAGCCGCTACGATCTGGCCGCGGTGCGTCAAGGTGACAGAGAGAGCTGGCGCGCA  
X 10 20 30 40 50 60  
CCAAATCCGCTCAGGTGACCAACCACTTCAACCGAGAGAGAGCTGCGCCACCATCGATGCTGTC  
|||||  
AGAACCCGCGCGACGACCTCGACCGCTGACCGAGAGAGAGAGCTGCGCCACCATCGATGCTGTC  
70 80 90 100 110 120 130 140

150 160 170 180 190 200 210 X  
CCCTTCAGAGAGCCGACAGACGATGACCCCGCGCGGTGAGCTGGCGGTGATGTGACGACAT  
|||||  
GCGTGCAGAGAGCCGACAGACGATGACCCCGCGCGGTGAGCTGGCGGTGATGTGACGACAT  
150 160 170 180 190 200 210

## 15. US-09-697-123b-21 (1-214)

US-09-697-123b-2 Sequence 20, Application US/09697123b

Initial Score = 109 Optimized Score = 159 Significance = -0.07  
Residue Identity = 79% Matches = 171 Mismatches = 34  
Gaps = 9 Conservative Substitutions = 0

X 10 20 30 40 50 60 70  
TCAGAGAGAGCCGCTACGATCTGGCCGCGGTGCGTCAAGGTGACAGAGAGCTGGCGCGCA  
|||||  
TCAGAGAGAGCCGCTACGATCTGGCCGCGGTGCGTCAAGGTGACAGAGAGCTGGCGCGCA  
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140  
CCAAATCCGCTCAGGTGACCAACCACTTCAACCGAGAGAGAGCTGCGCCACCATCGATGCTGTC  
|||||  
GCGAGCC-----GATCACCAGCTCGACCGCTGACCGAGAGAGAGCTGCGCCACCATCGATGCTGTC  
80 90 100 110 120 130

150 160 170 180 190 200 X  
GCGTGCAGAGAGCCGACAGACGATGACCCCGCGCGGTGAGCTGGCGGTGATGTGACGACAT  
|||||  
GCGTGCAGAGAGCCGACAGACGATGACCCCGCGCGGTGAGCTGGCGGTGATGTGACGACAT  
140 150 160 170 180 190 200 X

## 16. US-09-697-123b-21 (1-214)

US-09-697-123b-1 Sequence 16, Application US/09697123b

Initial Score = 109 Optimized Score = 162 Significance = -0.07  
Residue Identity = 78% Matches = 169 Mismatches = 39  
Gaps = 6 Conservative Substitutions = 0

X 10 20 30 40 50 60 70  
TCAGAGAGAGCCGCTACGATCTGGCCGCGGTGCGTCAAGGTGACAGAGAGCTGGCGCGCA  
|||||  
TCAGAGAGAGCCGCTACGATCTGGCCGCGGTGCGTCAAGGTGACAGAGAGCTGGCGCGCA  
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140  
CCAAATCCGCTCAGGTGACCAACCACTTCAACCGAGAGAGAGCTGCGCCACCATCGATGCTGTC  
|||||  
GCGAGCC-----GATCACCAGCTCGACCGCTGACCGAGAGAGAGCTGCGCCACCATCGATGCTGTC  
80 90 100 110 120 130

150 160 170 180 190 200 210 X  
GCGTGCAGAGAGCCGACAGACGATGACCCCGCGCGGTGAGCTGGCGGTGATGTGACGACAT  
|||||  
GCGTGCAGAGAGCCGACAGACGATGACCCCGCGCGGTGAGCTGGCGGTGATGTGACGACAT  
140 150 160 170 180 190 200 X

## 17. US-09-697-123b-21 (1-214)

US-09-697-123b-6 Sequence 6, Application US/09697123b

Initial Score = 109 Optimized Score = 162 Significance = -0.07  
Residue Identity = 78% Matches = 169 Mismatches = 39  
Gaps = 6 Conservative Substitutions = 0

X 10 20 30 40 50 60 70  
TCAGAGAGAGCCGCTACGATCTGGCCGCGGTGCGTCAAGGTGACAGAGAGAGCTGGCGCGCA  
|||||  
TCAGAGAGAGCCGCTACGATCTGGCCGCGGTGCGTCAAGGTGACAGAGAGAGCTGGCGCGCA  
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140  
CCAAATCCGCTCAGGTGACCAACCACTTCAACCGAGAGAGAGCTGCGCCACCATCGATGCTGTC  
|||||  
GCGAGCC-----GATCACCAGCTCGACCGCTGACCGAGAGAGAGCTGCGCCACCATCGATGCTGTC  
80 90 100 110 120 130

150 160 170 180 190 200 X  
GCGTGCAGAGAGCCGACAGACGATGACCCCGCGCGGTGAGCTGGCGGTGATGTGACGACAT  
|||||  
GCGTGCAGAGAGCCGACAGACGATGACCCCGCGCGGTGAGCTGGCGGTGATGTGACGACAT  
140 150 160 170 180 190 200 X

## 18. US-09-697-123b-21 (1-214)

US-09-697-123b-5 Sequence 5, Application US/09697123b

Initial Score = 109 Optimized Score = 162 Significance = -0.07  
Residue Identity = 78% Matches = 169 Mismatches = 39  
Gaps = 6 Conservative Substitutions = 0

X 10 20 30 40 50 60 70  
TCAGAGAGAGCCGCTACGATCTGGCCGCGGTGCGTCAAGGTGACAGAGAGCTGGCGCGCA  
|||||  
TCAGAGAGAGCCGCTACGATCTGGCCGCGGTGCGTCAAGGTGACAGAGAGCTGGCGCGCA  
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140  
CCAAATCCGCTCAGGTGACCAACCACTTCAACCGAGAGAGAGCTGCGCCACCATCGATGCTGTC  
|||||  
GCGAGCC-----GATCACCAGCTCGACCGCTGACCGAGAGAGAGCTGCGCCACCATCGATGCTGTC  
80 90 100 110 120 130

150 160 170 180 190 200 X  
GCGTGCAGAGAGCCGACAGACGATGACCCCGCGCGGTGAGCTGGCGGTGATGTGACGACAT  
|||||  
GCGTGCAGAGAGCCGACAGACGATGACCCCGCGCGGTGAGCTGGCGGTGATGTGACGACAT  
140 150 160 170 180 190 200 X

## 19. US-09-697-123b-21 (1-214)

US-09-697-123b-1 Sequence 1, Application US/09697123b

Initial Score = 109 Optimized Score = 170 Significance = -0.07  
Residue Identity = 82% Matches = 177 Mismatches = 31  
Gaps = 6 Conservative Substitutions = 0

X  
10 20 30 40 50 60 70  
TCACAGAGAGCGCTACGATCTGCCCCGGTGGTCAAGGTGAAGGTGGGCGGCA  
|||||  
TCACAGAGAGCGCTACGATCTGCCCCGGTGGTCAAGGTGAAGGTGGGCGGCA  
X 10 20 30 40 50 60 70  
CGAATCCGCGCTCAGGTGACACACACCGCTCAGGAGAGAGGTGGCGACATCGAGTACTGTGTC  
|||||  
CGAATCCGCGCTCAGGTGACACACACCGCTCAGGAGAGAGGTGGCGACATCGAGTACTGTGTC  
X 10 20 30 40 50 60 70  
GCGATCC-----GATCACACAGTTTCGACGCTGACCGAGAGAGATGTCGTCACCATCGAGTACTGTGTC  
80 90 100 110 120 130 140  
GCGATCC-----GATCACACAGTTTCGACGCTGACCGAGAGAGATGTCGTCACCATCGAGTACTGTGTC  
150 160 170 180 190 200 210  
GCGTGCACGAGGCGCCAGACCAAGATGACCGCCCGCGCGCTGCAAGGTGGCGGAGAT  
|||||  
GCGTGCACGAGGCGCCAGACCAAGATGACCGCCCGCGCGCTGCAAGGTGGCGGAGAT  
X 150 160 170 180 190 200 210  
GCGTGCACGAGGCGCCAGACCAAGATGACCGCCCGCGCGCTGCAAGGTGGCGGAGAT  
X 150 160 170 180 190 200 210

## 20. US-09-697-123B-21 (1-214)

US-09-697-123B-2 Sequence 22, Application US/09697123B

Initial Score = 108 Optimized Score = 161 Significance = -0.09  
Residue Identity = 78% Matches = 168 Mismatches = 40  
Gaps = 6 Conservative Substitutions = 0

X  
10 20 30 40 50 60 70  
TCACAGAGAGCGCTACGATCTGCCCCGGTGGTCAAGGTGAAGGTGGGCGGCA  
|||||  
TCACAGAGAGCGCTACGATCTGCCCCGGTGGTCAAGGTGAAGGTGGGCGGCA  
X 10 20 30 40 50 60 70  
GCGATCC-----GATCACACAGTTTCGACGCTGACCGAGAGAGATGTCGTCACCATCGAGTACTGTGTC  
80 90 100 110 120 130 140  
GCGATCC-----GATCACACAGTTTCGACGCTGACCGAGAGAGATGTCGTCACCATCGAGTACTGTGTC  
150 160 170 180 190 200 210  
GCGTGCACGAGGCGCCAGACCAAGATGACCGCCCGCGCGCTGCAAGGTGGCGGAGAT  
|||||  
GCGTGCACGAGGCGCCAGACCAAGATGACCGCCCGCGCGCTGCAAGGTGGCGGAGAT  
X 150 160 170 180 190 200 210  
GCGTGCACGAGGCGCCAGACCAAGATGACCGCCCGCGCGCTGCAAGGTGGCGGAGAT  
X 150 160 170 180 190 200 210

## 21. US-09-697-123B-21 (1-214)

US-09-697-123B-1 Sequence 12, Application US/09697123B

Initial Score = 106 Optimized Score = 158 Significance = -0.14  
Residue Identity = 79% Matches = 170 Mismatches = 37  
Gaps = 7 Conservative Substitutions = 0

X  
10 20 30 40 50 60 70  
TCACAGAGAGCGCTACGATCTGCCCCGGTGGTCAAGGTGAAGGTGGGCGGCA  
|||||  
TCACAGAGAGCGCTACGATCTGCCCCGGTGGTCAAGGTGAAGGTGGGCGGCA  
X 10 20 30 40 50 60 70  
GCGATCC-----GATCACACAGTTTCGACGCTGACCGAGAGAGATGTCGTCACCATCGAGTACTGTGTC  
80 90 100 110 120 130 140  
GCGATCC-----GATCACACAGTTTCGACGCTGACCGAGAGAGATGTCGTCACCATCGAGTACTGTGTC  
150 160 170 180 190 200 210  
GCGTGCACGAGGCGCCAGACCAAGATGACCGCCCGCGCGCTGCAAGGTGGCGGAGAT  
|||||  
GCGTGCACGAGGCGCCAGACCAAGATGACCGCCCGCGCGCTGCAAGGTGGCGGAGAT  
X 150 160 170 180 190 200 210  
GCGTGCACGAGGCGCCAGACCAAGATGACCGCCCGCGCGCTGCAAGGTGGCGGAGAT  
X 150 160 170 180 190 200 210

Initial Score = 83 Optimized Score = 152 Significance = -0.68  
Residue Identity = 73% Matches = 168 Mismatches = 40  
Gaps = 21 Conservative Substitutions = 0

X  
10 20 30 40 50 60 70  
TCACAGAGAGCGCTACGATCTGCCCCGGTGGTCAAGGTGAAGGTGGGCGGCA  
|||||  
TCACAGAGAGCGCTACGATCTGCCCCGGTGGTCAAGGTGAAGGTGGGCGGCA  
X 10 20 30 40 50 60 70  
GCGATCC-----GATCACACAGTTTCGACGCTGACCGAGAGAGATGTCGTCACCATCGAGTACTGTGTC  
80 90 100 110 120 130 140  
GCGATCC-----GATCACACAGTTTCGACGCTGACCGAGAGAGATGTCGTCACCATCGAGTACTGTGTC  
150 160 170 180 190 200 210  
GCGTGCACGAGGCGCCAGACCAAGATGACCGCCCGCGCGCTGCAAGGTGGCGGAGAT  
|||||  
GCGTGCACGAGGCGCCAGACCAAGATGACCGCCCGCGCGCTGCAAGGTGGCGGAGAT  
X 150 160 170 180 190 200 210  
GCGTGCACGAGGCGCCAGACCAAGATGACCGCCCGCGCGCTGCAAGGTGGCGGAGAT  
X 150 160 170 180 190 200 210

## 23. US-09-697-123B-21 (1-214)

US-09-697-123B-4 Sequence 4, Application US/09697123B

Initial Score = 82 Optimized Score = 157 Significance = -0.70  
Residue Identity = 78% Matches = 169 Mismatches = 38  
Gaps = 7 Conservative Substitutions = 0

X  
10 20 30 40 50 60 70  
TCACAGAGAGCGCTACGATCTGCCCCGGTGGTCAAGGTGAAGGTGGGCGGCA  
|||||  
TCACAGAGAGCGCTACGATCTGCCCCGGTGGTCAAGGTGAAGGTGGGCGGCA  
X 10 20 30 40 50 60 70  
GCGATCC-----GATCACACAGTTTCGACGCTGACCGAGAGAGATGTCGTCACCATCGAGTACTGTGTC  
80 90 100 110 120 130 140  
GCGATCC-----GATCACACAGTTTCGACGCTGACCGAGAGAGATGTCGTCACCATCGAGTACTGTGTC  
150 160 170 180 190 200 210  
GCGTGCACGAGGCGCCAGACCAAGATGACCGCCCGCGCGCTGCAAGGTGGCGGAGAT  
|||||  
GCGTGCACGAGGCGCCAGACCAAGATGACCGCCCGCGCGCTGCAAGGTGGCGGAGAT  
X 150 160 170 180 190 200 210  
GCGTGCACGAGGCGCCAGACCAAGATGACCGCCCGCGCGCTGCAAGGTGGCGGAGAT  
X 150 160 170 180 190 200 210

## 24. US-09-697-123B-21 (1-214)

US-09-697-123B-1 Sequence 13, Application US/09697123B

Initial Score = 81 Optimized Score = 155 Significance = -0.73  
Residue Identity = 76% Matches = 176 Mismatches = 32  
Gaps = 21 Conservative Substitutions = 0

X  
10 20 30 40 50 60 70  
TCACAGAGAGCGCTACGATCTGCCCCGGTGGTCAAGGTGAAGGTGGGCGGCA  
|||||  
TCACAGAGAGCGCTACGATCTGCCCCGGTGGTCAAGGTGAAGGTGGGCGGCA  
X 10 20 30 40 50 60 70  
GCGATCC-----GATCACACAGTTTCGACGCTGACCGAGAGAGATGTCGTCACCATCGAGTACTGTGTC  
80 90 100 110 120 130 140  
GCGATCC-----GATCACACAGTTTCGACGCTGACCGAGAGAGATGTCGTCACCATCGAGTACTGTGTC  
150 160 170 180 190 200 210  
GCGTGCACGAGGCGCCAGACCAAGATGACCGCCCGCGCGCTGCAAGGTGGCGGAGAT  
|||||  
GCGTGCACGAGGCGCCAGACCAAGATGACCGCCCGCGCGCTGCAAGGTGGCGGAGAT  
X 150 160 170 180 190 200 210  
GCGTGCACGAGGCGCCAGACCAAGATGACCGCCCGCGCGCTGCAAGGTGGCGGAGAT  
X 150 160 170 180 190 200 210

```

      80      90      100      110      120      130
150  GCCTGCAC-----GAGGGCAGACCAC---GATGACCGCCCGGCGGTGCGAGTGCCTGG
      |||||||
140  GCCTGCACGCGCTCTCAAGGTGCCAGGCCCGCCGTTATGACTGCTCCCGCGGCGGTGCGAGTGCCTGG
      150      160      170      180      190      200      210
210  ATGTGGACGACAT
      | |||||||
      AACCGACGACAT
220  X

```

25. US-09-697-123b-21 (1-214)  
US-09-697-123b-2 Sequence 25, Application US/09697123B

```

Initial Score      = 19   Optimized Score = 19   Significance = -2.18
Residue Identity = 100%  Matches          = 19   Mismatches   = 0
Gaps              = 0    Conservative Substitutions = 0
X
TCAAGCAGAAAGCGCTACGA
|||||
TCAAGCAGAAAGCGCTACGA
X
10

```

FastDB - Fast Pairwise Comparison of Sequences  
Release 5.4

Query sequence being compared:	US-09-697-123B-20 (1-205)
Number of sequences searched:	26
Number of scores above cutoff:	25

Results of the initial comparison of US-09-697-123B-20 (1-205) with:  
File : US09697123B.seq

PARAMETERS	
Similarity matrix	Unary
Mismatch penalty	1
Gap penalty	5.00
Gap size penalty	0.33
Cutoff score	12
Randomization group	0

Scores:	Mean	Median	Standard Deviation
---------	------	--------	--------------------

Times:	CPU	Total Elapsed
--------	-----	---------------

00:00:00.00

Number of residues: 5077

Number of scores above cutoff: 25

The scores below are sorted by initial score. Significance is calculated based on initial score.

A 100% identical sequence to the query sequence was found:

Sequence Name	Description	Length	Init. Score	Opt. Score	Sig. Frame
---------------	-------------	--------	-------------	------------	------------

Sequence Name	Description	Init. Opt.	Length	Score	Score	Sig.	Frame
---------------	-------------	------------	--------	-------	-------	------	-------

2.	US-09-697-123B-1	Sequence 15,	Application U	208	177	189	0	80	
3.	US-09-697-123B-1	Sequence 1,	Application U	208	172	184	0	69	0
4.	US-09-697-123B-3	Sequence 3,	Application U	208	170	182	0	65	0
5.	US-09-697-123B-2	Sequence 2,	Application U	208	168	177	0	60	0
6.	US-09-697-123B-1	Sequence 11,	Application U	208	166	177	0	56	0
7.	US-09-697-123B-8	Sequence 10,	Application U	208	165	176	0	54	0
8.	US-09-697-123B-8	Sequence 8,	Application U	208	163	173	0	49	0
9.	US-09-697-123B-1	Sequence 16,	Application U	208	161	171	0	45	0
10.	US-09-697-123B-6	Sequence 6,	Application U	208	161	171	0	45	0
11.	US-09-697-123B-5	Sequence 5,	Application U	208	161	171	0	45	0
12.	US-09-697-123B-2	Sequence 22,	Application U	208	160	170	0	42	0
13.	US-09-697-123B-1	Sequence 17,	Application U	208	160	170	0	42	0
14.	US-09-697-123B-1	Sequence 19,	Application U	208	159	176	0	40	0
15.	US-09-697-123B-2	Sequence 24,	Application U	208	155	166	0	31	0
16.	US-09-697-123B-4	Sequence 4,	Application U	207	150	178	0	20	0
17.	US-09-697-123B-2	Sequence 23,	Application U	208	147	171	0	13	0
18.	US-09-697-123B-1	Sequence 12,	Application U	207	144	172	0	0.07	0
19.	US-09-697-123B-9	Sequence 13,	Application U	223	138	160	-0.07	0	0
20.	US-09-697-123B-1	Sequence 9,	Application U	223	131	163	-0.22	0	0
21.	US-09-697-123B-1	Sequence 14,	Application U	214	122	162	-0.42	0	0
22.	US-09-697-123B-2	Sequence 7,	Application U	214	109	159	-0.71	0	0
23.	US-09-697-123B-7	Sequence 7,	Application U	214	109	158	-0.71	0	0
24.	US-09-697-123B-1	Sequence 18,	Application U	211	108	152	-0.74	0	0
**** 2 standard deviations below mean ****									
25.	US-09-697-123B-2	Sequence 25,	Application U	19	15	15	-2.81	0	0

1. US-09-697-123B-20 (1-205)  
US-09-697-123B-2 Sequence 20, Application US/09697123B

Initial Score	-	205	Optimized Score	-	205	Significance	-	1.433
Residue Identity	-	100%	Matches	-	205	Mismatches	-	0
Gaps	-	0	Conservative Substitutions	-			-	0

[illegible]

150 160 170 180 190 200  
ACGAGGCCACGCCACGAGTACCGTCCCGCATGCAGGTCCGGTGGAGACCGACGACAT  
ACGAGGCCACGCCACGAGTACCGTCCCGCATGCAGGTCCGGTGGAGACCGACGACAT  
ACGAGGCCACGCCACGAGTACCGTCCCGCATGCAGGTCCGGTGGAGACCGACGACAT  
150 160 170 180 190 200

2. US-09-697-123B-20 (1-205)  
US-09-697-123B-1 Sequence 15, Application US/09697123B

Initial Score	-	177	Optimized Score	-	189	Significance	-	0.80
Residue Identity	-	93%	Matches	-	195	Mismatches	-	10
Gaps	-	3	Conservative Substitutions	-			-	0

  

TCAAGGAGAACGCGTACGACTGCGCGCTGTCCGCCGCTTCAAGTCAACAAAGAACTTGCGCTGCACACGG	10	20	30	40	50	60	70
TCAAGGAGAACGCGTACGACTGCGCGCTGTCCGCCGCTTCAAGTCAACAAAGAACTTGCGCTGCACACGG	10	20	30	40	50	60	70
X	10	20	30	40	50	60	70

8. US-09-697-123B-20 (1-205)  
US-09-697-123B-8 Sequence 8, Application US/09697123B



Initial Score = 163 Optimized Score = 173 Significance = 0.49  
Residue Identity = 86% Matches = 179 Mismatches = 26  
Gaps = 3 Conservative Substitutions = 0

X  
TCAGAGAAACGGCTACGACCTGGCGCGCTGCGCCGCTACAGGTCAACAGAGCTGCGCTGACAGCGG  
TCAGAGAAACGGCTACGACCTGGCGCGCTGCGCCGCTACAGGTCAACAGAGCTGCGCTGACAGCGG  
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140  
GCCAGCCGATCACCAGCTGACGACCTGACGAGAAAGCTGCGCCGCTACAGGTCAACAGAGCTGCGCTGACAGCGG  
ATCATCGATCACCAGCTGACGACCTGACGAGAAAGCTGCGCCGCTACAGGTCAACAGAGCTGCGCTGACAGCGG  
X 10 20 30 40 50 60 70

150 160 170 180 190 200  
ACGAGGCGCCAGCCGCTACGACCTGACGAGAAAGCTGCGCCGCTACAGGTCAACAGAGCTGCGCTGACAGCGG  
ACGAGGCGCCAGCCGCTACGACCTGACGAGAAAGCTGCGCCGCTACAGGTCAACAGAGCTGCGCTGACAGCGG  
X 150 160 170 180 190 200

9. US-09-697-123b-20 (1-205)  
US-09-697-123b-1 Sequence 16, Application US/09697123b

Initial Score = 161 Optimized Score = 171 Significance = 0.45  
Residue Identity = 85% Matches = 177 Mismatches = 28  
Gaps = 3 Conservative Substitutions = 0

X  
TCAGAGAAACGGCTACGACCTGGCGCGCTGCGCCGCTACAGGTCAACAGAGCTGCGCTGACAGCGG  
TCAGAGAAACGGCTACGACCTGGCGCGCTGCGCCGCTACAGGTCAACAGAGCTGCGCTGACAGCGG  
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140  
GCCAGCCGATCACCAGCTGACGACCTGACGAGAAAGCTGCGCCGCTACAGGTCAACAGAGCTGCGCTGACAGCGG  
GCCAGCCGATCACCAGCTGACGACCTGACGAGAAAGCTGCGCCGCTACAGGTCAACAGAGCTGCGCTGACAGCGG  
X 80 90 100 110 120 130 140

150 160 170 180 190 200  
ACGAGGCGCCAGCCGCTACGACCTGACGAGAAAGCTGCGCCGCTACAGGTCAACAGAGCTGCGCTGACAGCGG  
ACGAGGCGCCAGCCGCTACGACCTGACGAGAAAGCTGCGCCGCTACAGGTCAACAGAGCTGCGCTGACAGCGG  
X 150 160 170 180 190 200

10. US-09-697-123b-20 (1-205)  
US-09-697-123b-6 Sequence 6, Application US/09697123b

Initial Score = 161 Optimized Score = 171 Significance = 0.45  
Residue Identity = 85% Matches = 177 Mismatches = 28  
Gaps = 3 Conservative Substitutions = 0

X  
TCAGAGAAACGGCTACGACCTGGCGCGCTGCGCCGCTACAGGTCAACAGAGCTGCGCTGACAGCGG  
TCAGAGAAACGGCTACGACCTGGCGCGCTGCGCCGCTACAGGTCAACAGAGCTGCGCTGACAGCGG  
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140  
GCCAGCCGATCACCAGCTGACGACCTGACGAGAAAGCTGCGCCGCTACAGGTCAACAGAGCTGCGCTGACAGCGG  
GCCAGCCGATCACCAGCTGACGACCTGACGAGAAAGCTGCGCCGCTACAGGTCAACAGAGCTGCGCTGACAGCGG  
X 80 90 100 110 120 130 140

150 160 170 180 190 200  
ACGAGGCGCCAGCCGCTACGACCTGACGAGAAAGCTGCGCCGCTACAGGTCAACAGAGCTGCGCTGACAGCGG  
ACGAGGCGCCAGCCGCTACGACCTGACGAGAAAGCTGCGCCGCTACAGGTCAACAGAGCTGCGCTGACAGCGG  
X 150 160 170 180 190 200

11. US-09-697-123b-20 (1-205)  
US-09-697-123b-5 Sequence 5, Application US/09697123b

Initial Score = 161 Optimized Score = 171 Significance = 0.45  
Residue Identity = 85% Matches = 177 Mismatches = 28  
Gaps = 3 Conservative Substitutions = 0

X  
TCAGAGAAACGGCTACGACCTGGCGCGCTGCGCCGCTACAGGTCAACAGAGCTGCGCTGACAGCGG  
TCAGAGAAACGGCTACGACCTGGCGCGCTGCGCCGCTACAGGTCAACAGAGCTGCGCTGACAGCGG  
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140  
GCCAGCCGATCACCAGCTGACGACCTGACGAGAAAGCTGCGCCGCTACAGGTCAACAGAGCTGCGCTGACAGCGG  
GCCAGCCGATCACCAGCTGACGACCTGACGAGAAAGCTGCGCCGCTACAGGTCAACAGAGCTGCGCTGACAGCGG  
X 80 90 100 110 120 130 140

150 160 170 180 190 200  
ACGAGGCGCCAGCCGCTACGACCTGACGAGAAAGCTGCGCCGCTACAGGTCAACAGAGCTGCGCTGACAGCGG  
ACGAGGCGCCAGCCGCTACGACCTGACGAGAAAGCTGCGCCGCTACAGGTCAACAGAGCTGCGCTGACAGCGG  
X 150 160 170 180 190 200

12. US-09-697-123b-20 (1-205)  
US-09-697-123b-2 Sequence 22, Application US/09697123b

Initial Score = 160 Optimized Score = 170 Significance = 0.42  
Residue Identity = 84% Matches = 176 Mismatches = 29  
Gaps = 3 Conservative Substitutions = 0

X  
TCAGAGAAACGGCTACGACCTGGCGCGCTGCGCCGCTACAGGTCAACAGAGCTGCGCTGACAGCGG  
TCAGAGAAACGGCTACGACCTGGCGCGCTGCGCCGCTACAGGTCAACAGAGCTGCGCTGACAGCGG  
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140  
GCCAGCCGATCACCAGCTGACGACCTGACGAGAAAGCTGCGCCGCTACAGGTCAACAGAGCTGCGCTGACAGCGG  
GCCAGCCGATCACCAGCTGACGACCTGACGAGAAAGCTGCGCCGCTACAGGTCAACAGAGCTGCGCTGACAGCGG  
X 80 90 100 110 120 130 140

150 160 170 180 190 200  
ACGAGGCGCCAGCCGCTACGACCTGACGAGAAAGCTGCGCCGCTACAGGTCAACAGAGCTGCGCTGACAGCGG  
ACGAGGCGCCAGCCGCTACGACCTGACGAGAAAGCTGCGCCGCTACAGGTCAACAGAGCTGCGCTGACAGCGG  
X 150 160 170 180 190 200

13. US-09-697-123b-20 (1-205)  
US-09-697-123b-1 Sequence 17, Application US/09697123b

Initial Score = 160 Optimized Score = 170 Significance = 0.42  
Residue Identity = 84% Matches = 176 Mismatches = 29  
Gaps = 3 Conservative Substitutions = 0

X  
TCAGAGAAACGGCTACGACCTGGCGCGCTGCGCCGCTACAGGTCAACAGAGCTGCGCTGACAGCGG  
TCAGAGAAACGGCTACGACCTGGCGCGCTGCGCCGCTACAGGTCAACAGAGCTGCGCTGACAGCGG  
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140  
GCCAGCCGATCACCAGCTGACGACCTGACGAGAAAGCTGCGCCGCTACAGGTCAACAGAGCTGCGCTGACAGCGG  
GCCAGCCGATCACCAGCTGACGACCTGACGAGAAAGCTGCGCCGCTACAGGTCAACAGAGCTGCGCTGACAGCGG  
X 80 90 100 110 120 130 140

5  
10  
20  
30  
40  
50  
60  
70  
CAAGGAAAGCGGTACGACCTGGCGGTGTCTGGCGGCTACAAAGTCAACACAGAAGCTCGGCTCGACGGG

```

|||||
TCAGAGAGAGCGCTACGACCTGGCCCGCTGCGCCCTACCAAGTCAACAGAACTGGCCCTGAACACCG
X 10 20 30 40 50 60 70
GGGACCCCATACCAAGCTCGACCTGACCCAGAGAAAGCTGTGCGCCACCATGAGTACCTGTGGCTTC
100 110 120 130 140
ATCATCTCGATACCAACACGACGCTGACCCGAAAGAGAGCTGTGCGCCACCATGAGTACCTGTGGCTTC
80 90 100 110 120 130 140
150
ACGAGGCG-----CAGCCCAAGATGACCTGCCCCG---ATCGAGGTGCCGTGGAGACCG
|||||
ACCAAGCTCTCAGAGTGGCCAGGCGCCCTTATGACTGTCCCGGGGGGTGAGGTGCGGGAACCG
150 160 170 180 190 200 210
200 X
ACGACAT
|||||
ACGACAT
220 X

```

20. US-09-697-123b-20 (1-205)  
US-09-697-123b-9 Sequence 9, Application US/09697123b

```

Initial Score = 131 Optimized Score = 163 Significance = -0.22
Residue Identity = 82% Matches = 183 Mismatches = 22
Gaps = 18 Conservative Substitutions = 0
X 10 20 30 40 50 60 70
TCAGAGAGAGCGCTACGACCTGGCCCGCTGCGCCCTTCAAGTCAACAGAACTGGCCCTGACACCGG
|||||
TCAAGAGAGAGCGCTACGACCTGGCCCGCTGCGCCCTTCAAGTCAACAGAACTGGCTGTGACACCGG
X 10 20 30 40 50 60 70
80 90 100 110 120 130 140
GCGAGCGGATCACCAGCTCGACGCTGACCCGAGAGAAAGCTGTGCGCCACCATGAGTACCTGTGGCTTC
|||||
GCGAGCGGATCACCAGCTCGACGCTGACCCGAGAGAAAGCTGTGCGCCACCATGAGTACCTGTGGCTTC
80 90 100 110 120 130 140
150 160 170 180 190
AC-----GAGGGCGAGCCAC---GATGACCTGCC---CGGATCTGAGGTGCCGCTGGAGACCG
|||||
ACACGCGCCGTAGCGATGAGCCAGCCGCGCTCATGACTGTCCCGGGGCGATCGAGTCCCGGTGAGAACCG
150 160 170 180 190 200 210
200 X
ACGACAT
|||||
ACGACAT
220 X

```

21. US-09-697-123b-20 (1-205)  
US-09-697-123b-1 Sequence 14, Application US/09697123b

```

Initial Score = 122 Optimized Score = 162 Significance = -0.42
Residue Identity = 81% Matches = 174 Mismatches = 31
Gaps = 9 Conservative Substitutions = 0
X 10 20 30 40 50 60 70
TCAGAGAGAGCGGCTACGACCTGGCCGCTGCGCCGCTTCAAGTCAACAGAACTGGCCCTGACACCGG
|||||
TCAAGAGAGAGCGGCTACGACCTGGCCGCTGCGCCGCTTCAAGTCAACAGAACTGGCCCTGACACCGG
X 10 20 30 40 50 60 70
80 90 100 110 120 130
GCGAGCGG-----ATCACCAGCTCGACGCTGACCCGAGAGAAAGCTGTGCGCCACCATGAGTACCTGTGC
|||||
CCGAGTGGCGGCTTACCGCTCGACGCTGACCCGAGAGCGGATGTCTGCGCCACCATGAGTACCTGTGC
80 90 100 110 120 130 140

```

```

140 150 160 170 180 190 200
GCGTTCACAGAGGCGCCAGCCACGATGACCGTCCC---CGGCATCGAGTGTCCCGGTGAGACCGACGACAT
|||||
GCGTTCACAGAGGCGCCAGCCACGATGACCGTCCC---CGGCATCGAGTGTCCCGGTGAGACCGACGACAT
150 160 170 180 190 200 210 X

```

22. US-09-697-123b-20 (1-205)  
US-09-697-123b-2 Sequence 21, Application US/09697123b

```

Initial Score = 109 Optimized Score = 159 Significance = -0.71
Residue Identity = 79% Matches = 171 Mismatches = 34
Gaps = 9 Conservative Substitutions = 0
X 10 20 30 40 50 60 70
TCAGAGAGAGCGGCTACGACCTGGCCGCTGCGCCGCTTCAAGTCAACAGAACTGGCCCTGACACCGG
|||||
TCAAGAGAGAGCGGCTACGACCTGGCCGCTGCGCCGCTTCAAGTCAACAGAACTGGCCCTGACACCGG
X 10 20 30 40 50 60 70

```

```

GCGAGCC-----GATCACCAGCTCGACGCTGACCCGAGAGAAAGCTGTGCGCCACCATGAGTACCTGTGC
|||||
CCAAATCCGGCTCAGGTGACCAACACGACCCCTCACCGAGAGAGCTGTGCCACCATGAGTACCTGTGC
80 90 100 110 120 130 140
140 150 160 170 180 190 200 X
GCGTTCACAGAGGCGCCAGCCACGATGACCGTCCC---CGGCATCGAGTGTCCCGGTGAGACCGACAT
|||||
GCGTTCACAGAGGCGCCAGCCACGATGACCGTCCC---CGGCATCGAGTGTCCCGGTGAGTGTGAGACAT
150 160 170 180 190 200 210 X

```

23. US-09-697-123b-20 (1-205)  
US-09-697-123b-7 Sequence 7, Application US/09697123b

```

Initial Score = 109 Optimized Score = 158 Significance = -0.71
Residue Identity = 79% Matches = 170 Mismatches = 35
Gaps = 9 Conservative Substitutions = 0
X 10 20 30 40 50 60 70
TCAGAGAGAGCGGCTACGACCTGGCCGCTGCGCCGCTTCAAGTCAACAGAACTGGCCCTGACACCGG
|||||
TCAAGAGAGAGCGGCTACGACCTGGCCGCTGCGCCGCTTCAAGTCAACAGAACTGGCTGTGCGCGTGC
X 10 20 30 40 50 60 70

```

```

GCGAGCC-----GATCACCAGCTCGACGCTGACCCGAGAGAAAGCTGTGCGCCACCATGAGTACCTGTGC
|||||
CCAAATCCGGCTTGTGACTGTGCCACGACGCTCACCGAGAGAGCTGTGCCACCATGAGTACCTGTGC
80 90 100 110 120 130 140
140 150 160 170 180 190 200 X
GCGTTCACAGAGGCGCCAGCCACGATGACCGTCCC---CGGCATCGAGTGTCCCGGTGAGACCGACAT
|||||
GCGTTCACAGAGGCGCCAGCCACGATGACCGTCCC---CGGCATCGAGTGTCCCGGTGAGTGTGAGACAT
150 160 170 180 190 200 210 X

```

24. US-09-697-123b-20 (1-205)  
US-09-697-123b-1 Sequence 18, Application US/09697123b

```

Initial Score = 108 Optimized Score = 152 Significance = -0.74
Residue Identity = 77% Matches = 163 Mismatches = 42
Gaps = 6 Conservative Substitutions = 0
X 10 20 30 40 50 60 70
TCAGAGAGAGCGGCTACGACCTGGCCGCTGCGCCGCTTCAAGTCAACAGAACTGGCCCTGACACCGG
|||||
TCAAGAGAGAGCGGCTACGACCTGGCCGCTGCGCCGCTTCAAGTCAACAGAACTGGCCCTGACACCGG
X 10 20 30 40 50 60 70

```







GCACAGCCATCATCTGTCGACTGTGACAGGAAAGACGCTGCCCAACATCAGTAACCTGGTGCGCCCTG  
80 90 100 110 120 130 140

GGGATCCGATTATCACCGATCCACGCTGACGACGAAGAAGACTGCTGCCCAACATCAGTAGTATGCTGGTGA  
80 90 100 110 120 130 140

X  
10 20 30 40 50 60 70  
TCAAGAGAGCGCTACGACCTGGCCCCGGCTGGCGCTACAGAAGTCAACAGAAGCTGGGCGCTGAACGCGC



|||||  
TCAGAGAGAGGCTTACGACCTGGCCCGCTGCGCCGTACAGGTCAAGAAAGCTGGGTGACAGCCGC  
X 10 20 30 40 50 60 70  
80 90 100 110 120 130 140  
GCCAGCCGATCTACCTGCTGACCTGTGACCGAGAAAGAGCTGTCGCCACCATGAGTACCTGGTGGCCCTGC  
|||  
GCCAGCCGATCTACCTGCTGACCTGTGACCGAGAAAGAGCTGTCGCCACCATGAGTACCTGGTGGCCCTGC  
|||  
GCCAGCCGATCTACCTGCTGACCTGTGACCGAGAAAGAGCTGTCGCCACCATGAGTACCTGGTGGCCCTGC  
80 90 100 110 120 130 140  
ACGAGAGGCGACA-----CCACGATGACCGCTCCCGGCGGTGAGAGTCCCGGTGAGGTGG  
|||  
ACGAGAGGCGCGTACGAGATGCGCCAGCCCGCTGATGACTGTCCCGGCGCATGAGAGTGGCGGTGAGACCG  
150 160 170 180 190 200 210  
X  
ACGACAT  
|||||  
ACGACAT  
220 X

20. US-09-697-123B-19 (1-208)  
US-09-697-123B-1 Sequence 13, Application US/09697123B

Initial Score = 143 Optimized Score = 169 Significance = -0.22  
Residue Identity = 82% Matches = 183 Mismatches = 25  
Gaps = 15 Conservative Substitutions = 0

X 10 20 30 40 50 60 70  
TCAGAGAGAGGCGCTACGACCTGGCCCGCTGAGCTCAAGTCAACAGAAAGCTGGGCTTGAACCCG  
|||||  
TCAGAGAGAGAGGCGCTACGACCTGGCCCGCTGAGCTCAAGTCAACAGAAAGCTGGGCTTGAACCCG  
X 10 20 30 40 50 60 70  
GCCAGCCGATCTACCTGCTGACCTGTGACCGAGAAAGAGCTGTCGCCACCATGAGTACCTGGTGGCCCTGC  
|||  
ATCATTCGATCTACCATCAGACCACTGACCGAGAAAGAGCTGTCGCCACCATGAGTACCTGGTGGCCCTGC  
80 90 100 110 120 130 140  
150 160 170 180 190 200 210  
AC-----GAGGCGCAGACAC--GATGACCGTCCCGGCGGTGAGAGTCCCGGTGAGAGTGG  
|||  
AC-----GAGGCGCAGACAC--GATGACCGTCCCGGCGGTGAGAGTCCCGGTGAGAGTGG  
|||  
ACGAGAGGCGCTACGAGCTGGCCCGCTGATGACTGTCCCGGCGGTGAGAGTCCCGGTGAGAAACCG  
150 160 170 180 190 200 210  
X  
ACGACAT  
|||||  
ACGACAT  
220 X

21. US-09-697-123B-19 (1-208)  
US-09-697-123B-1 Sequence 18, Application US/09697123B

Initial Score = 125 Optimized Score = 170 Significance = -0.59  
Residue Identity = 83% Matches = 176 Mismatches = 32  
Gaps = 3 Conservative Substitutions = 0

X 10 20 30 40 50 60 70  
TCAGAGAGAGGCGCTACGACCTGGCCCGCTGAGCTCAAGTCAACAGAAAGCTGGGCTTGAACCCG  
|||||  
TCAGAGAGAGGCGCTACGACCTGGCCCGCTGAGCTCAAGTCAACAGAAAGCTGGGCTTGAACCCG  
X 10 20 30 40 50 60 70  
GCCAGCCGATCTACCTGCTGACCTGTGACCGAGAAAGAGCTGTCGCCACCATGAGTACCTGGTGGCCCTGC  
|||  
GCCAGCCGATCTACCTGCTGACCTGTGACCGAGAAAGAGCTGTCGCCACCATGAGTACCTGGTGGCCCTGC  
80 90 100 110 120 130 140  
ACCGGCGCGACAGACCTACGACCTGGCCCGCTGAGCTCAAGTCAACAGAAAGCTGGGCTTGAACCCG  
|||  
ACCGGCGCGACAGACCTACGACCTGGCCCGCTGAGCTCAAGTCAACAGAAAGCTGGGCTTGAACCCG  
80 90 100 110 120 130 140

150 160 170 180 190 200  
TGCAGAGAGGCGCAGACCGATGACCGTCCCGGCGGTGAGAGTCCCGGTGAGAGTGGACAGCAT  
|||  
TGCATTCAGAGGCGCAGACCGATGACCGTCCCGGCGGTGAGAGTCCCGGTGAGAGTGGACAGCAT  
150 160 170 180 190 200 210

22. US-09-697-123B-19 (1-208)  
US-09-697-123B-2 Sequence 21, Application US/09697123B

Initial Score = 125 Optimized Score = 177 Significance = -0.59  
Residue Identity = 85% Matches = 184 Mismatches = 24  
Gaps = 6 Conservative Substitutions = 0

X 10 20 30 40 50 60 70  
TCAGAGAGAGGCGCTACGACCTGGCCCGCTGAGCTCAAGTCAACAGAAAGCTGGGCTTGAACCCG  
|||||  
TCAGAGAGAGGCGCTACGACCTGGCCCGCTGAGCTCAAGTCAACAGAAAGCTGGGCTTGAACCCG  
X 10 20 30 40 50 60 70  
GCCAGCC-----GATCAGCTGCTGACCTGTGACCGAGAAAGAGCTGTCGCCACCATGAGTACCTGGTGC  
|||  
CCAATTCGCGCTACAGTGAACCCAGCAGCAGCTTCAACAGAGAGAGCTGTGCTGACCATGAGTACCTGGTGC  
80 90 100 110 120 130 140

140 150 160 170 180 190 200  
GCCAGCCAGAGGCGCAGACCGATGACCGTCCCGGCGGTGAGAGTCCCGGTGAGAGTGGACAGCAT  
|||  
GCCAGCCAGAGGCGCAGACCGATGACCGTCCCGGCGGTGAGAGTCCCGGTGAGAGTGGACAGCAT  
150 160 170 180 190 200 210 X

23. US-09-697-123B-19 (1-208)  
US-09-697-123B-1 Sequence 14, Application US/09697123B

Initial Score = 116 Optimized Score = 168 Significance = -0.78  
Residue Identity = 81% Matches = 175 Mismatches = 33  
Gaps = 6 Conservative Substitutions = 0

X 10 20 30 40 50 60 70  
TCAGAGAGAGGCGCTACGACCTGGCCCGCTGAGCTCAAGTCAACAGAAAGCTGGGCTTGAACCCG  
|||||  
TCAGAGAGAGGCGCTACGACCTGGCCCGCTGAGCTCAAGTCAACAGAAAGCTGGGCTTGAACCCG  
X 10 20 30 40 50 60 70  
GCCAGCCG-----ATCAGCTGCTGACCTGTGACCGAGAAAGAGCTGTCGCCACCATGAGTACCTGGTGC  
|||  
CCGAGTGGCGCGTACCGCGCTGACCGACCGCTACCGAGAGGATGCTGCTGCGCACCATGAGTACCTGGTGC  
80 90 100 110 120 130 140

140 150 160 170 180 190 200  
GCCAGCCAGAGGCGCAGACCGATGACCGTCCCGGCGGTGAGAGTCCCGGTGAGAGTGGACAGCAT  
|||  
GCCAGCCAGAGGCGCAGACCGATGACCGTCCCGGCGGTGAGAGTCCCGGTGAGAGTGGACAGCAT  
150 160 170 180 190 200 210 X

24. US-09-697-123B-19 (1-208)  
US-09-697-123B-7 Sequence 7, Application US/09697123B

Initial Score = 115 Optimized Score = 177 Significance = -0.80  
Residue Identity = 85% Matches = 184 Mismatches = 24  
Gaps = 6 Conservative Substitutions = 0

X 10 20 30 40 50 60 70  
TCAGAGAGAGGCGCTACGACCTGGCCCGCTGAGCTCAAGTCAACAGAAAGCTGGGCTTGAACCCG  
|||||  
TCAGAGAGAGGCGCTACGACCTGGCCCGCTGAGCTCAAGTCAACAGAAAGCTGGGCTTGAACCCG  
X 10 20 30 40 50 60 70  
GCCAGCCG-----ATCAGCTGCTGACCTGTGACCGAGAAAGAGCTGTCGCCACCATGAGTACCTGGTGC  
|||  
CCGAGTGGCGCGTACCGCGCTGACCGACCGCTACCGAGAGGATGCTGCTGCGCACCATGAGTACCTGGTGC  
80 90 100 110 120 130 140





8. US-09-697-123B-18 (1-211)  
US-09-697-123B-8 Sequence 8, Application US/09697123B

Initial Score =	117	Optimized Score =	164	Significance =	0.20
Residue Identity =	80%	Matches =	170	Mismatches =	38
Gaps =	3	Conservative Substitutions			0

TCAAAGGAAAGCGCTACACATGGCCGCCGTGGTGGATCAAGTTCAAAGAAGCTGGCGCATCACCCAGAG  
X 10 20 30 40 50 60 70

TCAAAGGAAAGCGCTACACATGGCCGCCGTGGTGGATCAAGTTCAAAGAAGCTGGCGCATCACCCAGAG  
X 10 20 30 40 50 60 70

ACCCGGCCGACACGACCTTCGACACAGCTGACCGAAGAGACTCTGCCACATCCAGTACCTGGTGGGG  
 80 90 100 110 120 130 140  
 AATCATCC-----GATCACACGACGACGCTGACCCGACAAAGACACTCTGCCACATCCAGTATCTGGTGGGG  
 80 90 100 110 120 130 140

150 160 170 180 190 200 210  
 TGCATCAGGGCGACCAAGACGATGCACCGTCCCGGTGAGAGTGCAGTGCAGCAGCAGT  
 150 160 170 180 190 200 X  
 TTGACGAGGGCCAGCCACGAGTGCACCGTCCCGGGGGGTGCAGTGCCTGGAACCGACGACAT

9. US-09-697-123B-18 (1-211)  
US-09-697-123B-2 Sequence 2, Application US/09697123B

Initial score	=	117	Optimized score	=	161	Significance	=	0.20
Residue Identity	=	79%	Matches	=	167	Mismatches	=	41
Gaps	=	3	Conservative Substitutions	=	0			

X 10 20 30 40 50 60 70  
 TCAAGAGAAAGCGCTACAGACTGCGCCGGTGCGGTCAAGGTAAACAAGACTTGGCATCAACGAGAA  
 TCAAGAGAAAGCGCTACAGACTGCGCCGGGTGGCGGCTCAAGGTCAACAAGACTGCGGTCAAGCTGG  
 X 10 20 30 40 50 60 70

ACCCGCGCCACAGCAGCTGCACACGCTGACCAAGAAGAGAGCTGTGCCACCATCAGTACTGGTG  
80      90      100      110      120      130      140  
          | | | | | | | | | | | | | | | | | | | | | |  
GCAGCGCATCACACAGCTG--ACGTCGACCGAAGAAAGCTGTGACCATCAGTACTGGTGCG  
90      100      110      120      130      14

150 160 170 180 190 200 210  
 TGCATAGGCGCAGACAGATAGCTCCCGGTGAGTCCAGTCCCTCCGTCAGATCAGACATC  
 150 160 170 180 190 200 210  
 TGCACGAGGTCAGTCGCGATGACGATTCGCCGGGGCGGCGAGATCCCGTGGAGGCCACGACAT

10. US-09-697-123B-18 (1-211)  
US-09-697-123B-1 Sequence 13, Application US/09697123B

Initial Score =	116	Optimized Score =	152	Significance =	0.17
Residue Identity =	73%	Matches =	167	Mismatches =	41
Gaps =	18	Conservative Substitutions			0

X 10 20 30 40 50 60 70  
 TTTAAGGAAACCGCTACGACCTGGCCCGGCGTGTACAAAGTCAAGAAAGCTGGGCATACCCGAGA  
 TTTAAGGAAACCGCTACGACCTGGCCCGGCGTGTACAAAGTCAAGAAAGCTGGGCCTAACCCGACCG  
 TTTAAGGAAACCGCTACGACCTGGCCCGGCGTGTACAAAGTCAAGAAAGCTGGGCCTAACCCGACCG  
 TTTAAGGAAACCGCTACGACCTGGCCCGGCGTGTACAAAGTCAAGAAAGCTGGGCCTAACCCGACCG

80 90 100 110 120 130 140  
 ACCGGCGGACACACATCTGACACCACTGACGAGAGAGCTGTGGACCATGCTACCTGGTGGC  
 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111  
 ATCATTCGATCACCAC--ACACGCTGACGGAAGAGCTGTGCTCCACCATGTGACTCTGGTCC  
 80 90 100 110 120 130 140

TGCATCAGAGGC-----GAC AAGAGATGACCGCTCCGGTGGAGTCAGAGCGCCCTCAGAG  
 150 160 170 180 190 200  
 TTCACACAGCCCTTCAGAGGTTGGCCAGAGGCCCGCTTATGACACTCCCGCGGGCTCAGAGTTCGGCTGCAAA  
 150 160 170 180 190 200 210

210  
TCGACGACAT  
|||||  
CCGACGACAT  
220 X

11. US-09-697-123B-18 (1-211)  
US-09-697-123B-9 Sequence 9, Application US/09697123B

Initial Score	-	116	Optimized Score	-	150	Significance	-	0.17
Residue Identity	-	73%	Matches	-	165	Mismatches	-	433
Gaps	-	18	Conservative Substitutions	-			-	0

X 10 20 30 40 50 60 70  
 TCAAGAGAAAGCGGTACACAGCTGGCCCGCTGGTGGTCAAGGTCAAGAAAGTGGCGATCACCGAAG  
 TCAAGGAAAGCGGTACACAGCTGGCCCGCTGGTGGTCAAGGTCAAGAAAGTGGCGATCACCGAAG  
 TCAAGGAAAGCGGTACACAGCTGGCCCGCTGGTGGTCAAGGTCAAGAAAGTGGCGATCACCGAAG  
 X 10 20 30 40 50 60 70

ACCGGGCGACAGCTGCGACCGAGCGAGCGTCTGCGACCGATTCGACTGTGTGGCGG  
 80 90 100 110 120 130 140  
 TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT  
 1  
 GCGAGCGATCAGC---TCGTCACGCTGACCGAGAGACGCTGCGCCACCTTCATTAACCTGTCGGG  
 80 90 100 110 120 130 140

TGCATAGGGCCACGAG-----ACGATGACCTCCCGGAGTGTGAGTCCCGCTGAGAG  
150 160 170 180 190 200  
TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT  
TGACACCACCCGTCAGGATGCCAGCCCCCTCATCTCTCTCCCAGCAGATGAGTCCCGTAGAA  
150 160 170 180 190 200 210

210  
TCGACGACAT  
|||||  
CCGACGACAT  
220 X

12. US-09-697-123B-18 (1-211)  
US-09-697-123B-1 Sequence 14, Application US/09697123B

Initial Score	=	115	Optimized Score	=	166	Significance	=	0.14
Residue Identity	=	80%	Matches	=	172	Mismatches	=	39
Gaps	=	3	Conservative Substitutions	=			=	0

TCAGGAGACGCGTACGACCTGCCCCGGTGCGGTCAAGGTCACAAACAACTGGCATACCGAG-  
 10 20 30 40 50 60 70  
 TCAGGAGACGCGTACGACCTGCCCCGGTGCGGTCAAGGTCACAAACAACTGGCATACCGAG-  
 10 20 30 40 50 60 70  
 TCAGGAGACGCGTACGACCTGCCCCGGTGCGGTCAAGGTCACAAACAACTGGCATACCGAG-  
 10 20 30 40 50 60 70

[illegible]

150 160 170 180 190 200 210  
 GCGTGCATAGGCGCAAGAAGATGACCGTCCGGGAGTGCAGCAGTCCCGTGGAGGTGCACACATCT  
 150 160 170 180 190 200 210  
 GCGTGCATAGGCGCAAGAAGATGACCGTCCGGGAGTGCAGCAGTCCCGTGGAGGTGCACACATCT  
 150 160 170 180 190 200 210  
 GCGTGCATAGGCGCAAGAAGATGACCGTCCGGGAGTGCAGCAGTCCCGTGGAGGTGCACACATCT

13. US-09-697-123B-18 (1-211)  
US-09-697-123B-2 Sequence 22, Application US/09697123B

Optimized Score	=	158	Significance	=	0.11
Matches	=	164	Mismatches	=	44
Conservative Substitutions	=	0			

150 160 170 180 190 200 210  
 TGCATCAGGGGACCAAGCATGACCCGTCGAGTCGAGGTCGACGACATCAT  
 TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT  
 TGCGAGAGGGTCAGACCAACGATGACCTTTCCGGGGGGCTCGAGGTGCGGTGAAACCGACACAT

70 80 90 100 110 120 130 140  
AGACCCGCCACACGACCTTCGACACGCTGACCGAAGGAGACGTCGTGCGCACCATTGACGTACCTGTGCG  
11 1111 1111 1111 1111 1111 1111 1111  
CCAAATCCGGCTCAGGTGACCAACACACACCTCACCCGAGGAGAACGTCGTGCGCACCATCGAATACCTGTGCGC

TTAGGAGAACCGCTACGACTGGCCCGGGTGGGTGCGTACAAAGTCAACAAGAGCTGGGCATCCGACAG  
 |||||  
 TTAGGAGAACCGCTACGACTGG-TCGGGTGGCGCGTTACAAAGTCAACAAAAGCTGGTCTGAACGTGCG  
 |||||  
 X 10 20 30 40 50 60 70

25. US-09-697-123B-18 (1-211)  
US-09-697-123B-2 Sequence 25, Application US/09697123B

Initial Score	=	19	Optimized Score	=	19	Significance	=	-2.6
Residue Identity	=	100%	Matches	=	19	Mismatches	=	0
Gaps	=	0	Conservative Substitutions	=			=	0
<pre>       X      10      X TCACGACGAAACGCTACGA       TCACGACGAAACGCTACGA       X      10      X </pre>								





```

      80      90      100      110      120      130      140
CGTCCCGGATCACAGACACCACTCTGACCGAAGAGGACGTGTCGCCACCATCGAGTACCTGGTCCGCTGC
|||||
ATCATCGGATCACACACACGACGCTACCGCAAGAAGACGTCGTGCCACCATCGAGTATCTGGTCCGCTGC
      80      90      100      110      120      130      140
      150      160      170      180      190      200      210
ACGAGGGCCACACCACTACCGTCCCGGGCGAGTCGAGGTGCGGGTGGAAACCGAGACAT
|||||
ACGAGGGCCAGGCGCACCATGACCGTCCCGGGCGGGTTCGAGGTGCGGGTGGAAACCGAGACAT
      150      160      170      180      190      200      210
X
```

3. US-09-697-123B-17 (1-208)  
US-09-697-123B-1 Sequence 19, Application US/09697123B

Initial Score = 183 Optimized Score = 183 Significance = 0.64  
Residue Identity = 87% Matches = 183 Mismatches = 25  
Gaps = 0 Conservative Substitutions = 0

```

      10      20      30      40      50      60      70
TCAAGGAGAGCGCTACGACCTCGCGGGTGGCGGCTACAAGGTCAACAAGAGCTCGGCTGAACACCG
|||||
TCAAGGAGAGCGCTACGACCTCGCGGGTGGCGGCTACAAGGTCAACAAGAGCTCGGCTGAACACCG
      10      20      30      40      50      60      70
X
      80      90      100      110      120      130      140
CGTCCCGGATCACAGACACCACTCTACCGAAGAGGACGTGTCGCCACCATCGAGTACCTGGTCCGCTGC
|||||
GCCAGCCGATCACGTGTCGACCTCTACCGAGAGAGACGTCGTGCCACCATCGAGTACCTGGTCCGCTGC
      80      90      100      110      120      130      140
X
```

```

      150      160      170      180      190      200      210
ACGAGGGCCACACCACTGACCGTCCCGGGCGGAGTCGAGGTGCGGGTGGAAACCGAGACAT
|||||
ACGAGGGCCAGACGATGACCGTCCCGGGCGGTCGAGGTGCGGGTGGAAACCGAGACAT
      150      160      170      180      190      200      210
X
```

4. US-09-697-123B-17 (1-208)  
US-09-697-123B-1 Sequence 15, Application US/09697123B

Initial Score = 183 Optimized Score = 183 Significance = 0.64  
Residue Identity = 87% Matches = 183 Mismatches = 25  
Gaps = 0 Conservative Substitutions = 0

```

      10      20      30      40      50      60      70
TCAAGGAGAGCGCTACGACCTCGCGGGTGGCGGCTACAAGGTCAACAAGAGCTCGGCTGAACACCG
|||||
TCAAGGAGAGCGCTACGACCTCGCGGGTGGCGGCTACAAGGTCAACAAGAGCTCGGCTGAACACCG
      10      20      30      40      50      60      70
X
      80      90      100      110      120      130      140
CGTCCCGGATCACAGACACCACTCTGACCGAAGAGGACGTGTCGCCACCATCGAGTACCTGGTCCGCTGC
|||||
GTGAGCGGATCACCACTGACGCTGACCGAGGAGAGCTGTCGCCACCATCGAGTACCTGGTCCGCTGC
      80      90      100      110      120      130      140
X
```

```

      150      160      170      180      190      200      210
ACGAGGGCCACACCACTGACCGTCCCGGGCGGAGTCGAGGTGCGGGTGGAAACCGAGACAT
|||||
ACGAGGGTCAGGCCACGATGACCGTCCCGGGCGGATCGAGGTGCGGGTGGAAACCGAGACAT
      150      160      170      180      190      200      210
X
```

5. US-09-697-123B-17 (1-208)  
US-09-697-123B-1 Sequence 1, Application US/09697123B

Initial Score = 182 Optimized Score = 182 Significance = 0.62  
Residue Identity = 87% Matches = 182 Mismatches = 26  
Gaps = 0 Conservative Substitutions = 0

X 10 20 30 40 50 60 70

```

      80      90      100      110      120      130      140
TCAAGGAGAGCGCTACGACCTCGCGGGTGGCGGCTACAAGGTCAACAAGAGCTCGGCTGAACACCG
|||||
TCAAGGAGAGCGCTACGACCTCGCGGGTGGCGGCTACAAGGTCAACAAGAGCTCGGCTGAACACCG
      10      20      30      40      50      60      70
X
      80      90      100      110      120      130      140
CGTCCCGGATCACAGACACCACTCTGACCGAAGAGGACGTGTCGCCACCATCGAGTACCTGGTCCGCTGC
|||||
GCGATCGGATCACCACTGTCACGCTGACCGAGAGAGACGTCGTGCCACCATCGAGTACCTGGTCCGCTGC
      80      90      100      110      120      130      140
X
```

```

      150      160      170      180      190      200      210
ACGAGGGCCACACCACTGACCGTCCCGGGCGGAGTCGAGGTGCGGGTGGAAACCGAGACAT
|||||
ACGAGGGCCAGCAGCAGATGACCGTCCCGGGCGGACCGAGGTGCGGGTGGAGACCGAGACAT
      150      160      170      180      190      200      210
X
```

6. US-09-697-123B-17 (1-208)  
US-09-697-123B-2 Sequence 24, Application US/09697123B

Initial Score = 181 Optimized Score = 181 Significance = 0.60  
Residue Identity = 87% Matches = 181 Mismatches = 27  
Gaps = 0 Conservative Substitutions = 0

```

      10      20      30      40      50      60      70
TCAAGGAGAGCGCTACGACCTCGCGGGTGGCGGCTACAAGGTCAACAAGAGCTCGGCTGAACACCG
|||||
TCAAGGAGAGCGCTACGACCTCGCGGGTGGCGGCTACAAGGTCAACAAGAGCTCGGCTGAACACCG
      10      20      30      40      50      60      70
X
```

```

      80      90      100      110      120      130      140
CGTCCCGGATCACAGACACCACTCTGACCGAAGAGGACGTGTCGCCACCATCGAGTACCTGGTCCGCTGC
|||||
AGAATCGGCAACCACTGACCGTCCCGGGTGGCGGCTACAAGGTCAACAAGAGCTCGGCTGAACACCG
      80      90      100      110      120      130      140
X
```

```

      150      160      170      180      190      200      210
ACGAGGGCCACACCACTGACCGTCCCGGGCGGAGTCGAGGTGCGGGTGGAAACCGAGACAT
|||||
ACGAGGGCCACCGCAGATGAGGTGCGGGTGGCGGTCGAGGTGCGGGTGGAGACCGAGACAT
      150      160      170      180      190      200      210
X
```

7. US-09-697-123B-17 (1-208)  
US-09-697-123B-1 Sequence 11, Application US/09697123B

Initial Score = 181 Optimized Score = 181 Significance = 0.60  
Residue Identity = 87% Matches = 181 Mismatches = 27  
Gaps = 0 Conservative Substitutions = 0

```

      10      20      30      40      50      60      70
TCAAGGAGAGCGCTACGACCTCGCGGGTGGCGGCTACAAGGTCAACAAGAGCTCGGCTGAACACCG
|||||
TCAAGGAGAGCGCTACGACCTCGCGGGTGGCGGCTACAAGGTCAACAAGAGCTCGGCTGAACACCG
      10      20      30      40      50      60      70
X
```

```

      80      90      100      110      120      130      140
CGTCCCGGATCACAGACACCACTCTGACCGAAGAGGACGTGTCGCCACCATCGAGTACCTGGTCCGCTGC
|||||
GCCAGCCATCACCACTGACGCTGACCGAGGAGACGTCGTGCCACCATCGAATACCTGGTCCGCTGC
      80      90      100      110      120      130      140
X
```

```

      150      160      170      180      190      200      210
ACGAGGGCCACACCACTGACCGTCCCGGGCGGAGTCGAGGTGCGGGTGGAAACCGAGACAT
|||||
ACGAGGGCCAGCAGATGACCGTCCCGGGCGGTCGAGGTGCGGGTGGAGACCGAGACAT
      150      160      170      180      190      200      210
X
```

8. US-09-697-123B-17 (1-208)  
US-09-697-123B-1 Sequence 10, Application US/09697123B

Initial Score = 179 Optimized Score = 179 Significance = 0.56  
Residue Identity = 86% Matches = 179 Mismatches = 29  
Gaps = 0 Conservative Substitutions = 0

X 10 20 30 40 50 60 70  
TCAAGGAGAGCGCTACGACCTCGCGCGGTGGCGCTACAAGGTCAACAAGTCTGGCTGAACACCG  
|||||  
TCAAGGAGAGCGCTACGACCTCGCGCGGTGGCGCTACAAGGTCAACAAGTCTGGCTGAACACCG  
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140  
CGTCCCGCATCAGACACCACTCTGACCGAGAGACGCTCGCGCACCATCGACTACCTGGTCCGCTGC  
|||||  
GCCAGCCCATCACCAGCTCGACGCTGACCGAGAGACGCTCGCGCACCATCGACTACCTGGTCCGCTGC  
80 90 100 110 120 130 140

150 160 170 180 190 200 X  
ACGAGGCCACACCACTGACCGTCCCGGGCGAGTTCGAGGTGCGGGTGAACCGACGACAT  
|||||  
ACGAGGCCACACCACTGACCGTCCCGGGCGAGTTCGAGGTGCGGGTGAACCGACGACAT  
150 160 170 180 190 200 X

9. US-09-697-123B-17 (1-208)  
US-09-697-123B-3 Sequence 3, Application US/09697123B

Initial Score = 179 Optimized Score = 179 Significance = 0.56  
Residue Identity = 86% Matches = 179 Mismatches = 29  
Gaps = 0 Conservative Substitutions = 0

X 10 20 30 40 50 60 70  
TCAAGGAGAGCGCTACGACCTCGCGCGGTGGCGCTACAAGGTCAACAAGTCTGGCTGAACACCG  
|||||  
TCAAGGAGAGCGCTACGACCTCGCGCGGTGGCGCTACAAGGTCAACAAGTCTGGCTGAACACCG  
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140  
CGTCCCGCATCAGACACCACTCTGACCGAGAGACGCTCGCGCACCATCGACTACCTGGTCCGCTGC  
|||||  
GCGATCCGATCACCAGCTCCACGCTGACCGAGAGAGAGCTCGTCCGCGCACCATCGACTACCTGGTCCGCTGC  
80 90 100 110 120 130 140

150 160 170 180 190 200 X  
ACGAGGCCACACCACTGACCGTCCCGGGCGAGTTCGAGGTGCGGGTGAACCGACGACAT  
|||||  
ACGAGGTTCAGACACGATGACCGTTCGCGGGCGACCGAGGTTCGCGTGGAGACCGACGACAT  
150 160 170 180 190 200 X

10. US-09-697-123B-17 (1-208)  
US-09-697-123B-1 Sequence 16, Application US/09697123B

Initial Score = 178 Optimized Score = 178 Significance = 0.53  
Residue Identity = 85% Matches = 178 Mismatches = 30  
Gaps = 0 Conservative Substitutions = 0

X 10 20 30 40 50 60 70  
TCAAGGAGAGCGCTACGACCTCGCGCGGTGGCGCTACAAGGTCAACAAGTCTGGCTGAACACCG  
|||||  
TCAAGGAGAGCGCTACGACCTCGCGCGGTGGCGCTACAAGGTCAACAAGTCTGGCTGAACACCG  
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140  
CGTCCCGCATCAGACACCACTCTGACCGAGAGACGCTCGCGCACCATCGACTACCTGGTCCGCTGC  
|||||  
GCGAGCCCATCAGCTCGACGCTGACCGAGAGAGAGCTCGTGGCGCACCATCGAATATCTGGTCCGCTGC  
80 90 100 110 120 130 140

150 160 170 180 190 200 X  
ACGAGGCCACACCACTGACCGTCCCGGGCGAGTTCGAGGTGCGGGTGAACCGACGACAT  
|||||  
ACGAGGTTCAGACACGATGACCGTTCGCGGGCGGCTCGAGGTGCGGGTGAACCGACGACAT  
150 160 170 180 190 200 X

11. US-09-697-123B-17 (1-208)  
US-09-697-123B-6 Sequence 6, Application US/09697123B

Initial Score = 178 Optimized Score = 178 Significance = 0.53  
Residue Identity = 85% Matches = 178 Mismatches = 30  
Gaps = 0 Conservative Substitutions = 0

X 10 20 30 40 50 60 70  
TCAAGGAGAGCGCTACGACCTCGCGCGGTGGCGCTACAAGGTCAACAAGTCTGGCTGAACACCG  
|||||  
TCAAGGAGAGCGCTACGACCTCGCGCGGTGGCGCTACAAGGTCAACAAGTCTGGCTGAACACCG  
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140  
CGTCCCGCATCAGACACCACTCTGACCGAGAGAGCTCGTCCGCGCACCATCGACTACCTGGTCCGCTGC  
|||||  
GCGAGCCCATCAGCTCGTGGAGCGCTGACCGAGAGAGAGCTCGTCCGCGCACCATCGAATATCTGGTCCGCTGC  
80 90 100 110 120 130 140

150 160 170 180 190 200 X  
ACGAGGCCACACCACTGACCGTCCCGGGCGAGTTCGAGGTGCGGGTGAACCGACGACAT  
|||||  
ACGAGGTTCAGACACGATGACCGTTCGCGGGCGGCTCGAGGTGCGGGTGAACCGACGACAT  
150 160 170 180 190 200 X

12. US-09-697-123B-17 (1-208)  
US-09-697-123B-5 Sequence 5, Application US/09697123B

Initial Score = 178 Optimized Score = 178 Significance = 0.53  
Residue Identity = 85% Matches = 178 Mismatches = 30  
Gaps = 0 Conservative Substitutions = 0

X 10 20 30 40 50 60 70  
TCAAGGAGAGCGCTACGACCTCGCGCGGTGGCGCTACAAGGTCAACAAGTCTGGCTGAACACCG  
|||||  
TCAAGGAGAGCGCTACGACCTCGCGCGGTGGCGCTACAAGGTCAACAAGTCTGGCTGAACACCG  
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140  
CGTCCCGCATCAGACACCACTCTGACCGAGAGAGCTCGTCCGCGCACCATCGACTACCTGGTCCGCTGC  
|||||  
GCGAGCCCATCAGCTCGTGGAGCGCTGACCGAGAGAGAGCTCGTCCGCGCACCATCGAATATCTGGTCCGCTGC  
80 90 100 110 120 130 140

150 160 170 180 190 200 X  
ACGAGGCCACACCACTGACCGTCCCGGGCGAGTTCGAGGTGCGGGTGAACCGACGACAT  
|||||  
ACGAGGTTCAGACACGATGACCGTTCGCGGGCGGCTCGAGGTGCGGGTGAACCGACGACAT  
150 160 170 180 190 200 X

13. US-09-697-123B-17 (1-208)  
US-09-697-123B-2 Sequence 22, Application US/09697123B

Initial Score = 177 Optimized Score = 177 Significance = 0.51  
Residue Identity = 85% Matches = 177 Mismatches = 31  
Gaps = 0 Conservative Substitutions = 0

X 10 20 30 40 50 60 70  
TCAAGGAGAGCGCTACGACCTCGCGCGGTGGCGCTACAAGGTCAACAAGTCTGGCTGAACACCG  
|||||  
TCAAGGAGAGCGCTACGACCTCGCGCGGTGGCGCTACAAGGTCAACAAGTCTGGCTGAACACCG  
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140  
CGTCCCGCATCAGACACCACTCTGACCGAGAGAGCTCGTCCGCGCACCATCGACTACCTGGTCCGCTGC  
|||||  
GCGAGCCCATCAGCTCGTGGAGCGCTGACCGAGAGAGAGCTCGTCCGCGCACCATCGAATATCTGGTCCGCTGC  
80 90 100 110 120 130 140



```
|||||
TCAAGGAGAGCGCTACGACCTGGCCCGCTCGCCGCTACAAGGTCAACAAGAGCTGGCGCTGAACACCG
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140
CGTCCCGGATCAGGAGGACACTCTGACCGAAGAGGACGTCGTGCGCACCATCGAGTACTGTCGCGCTGC
|||||
ATCATCGGATCACCAACGACGCTGACCGAAGAGACGTCGTGCGCACCATCGAGTACTGTCGCGCTGC
80 90 100 110 120 130 140

150 160 170 180 190 200 210
ACGAGGCG-----CACACCACGATGACCGTCCCGCGGAGTCGAGGTCCCGGTGGAAACCG
|||||
ACACGCTCTCAGGCTGGCGAGCGCCCGCTTATGACTGTCCCGCGGGTTCGAGGTCCCGGTGGAAACCG
150 160 170 180 190 200 210

X
ACGACAT
ACGACAT
220 X
```

## 20. US-09-697-123B-17 (1-208)

US-09-697-123B-7 Sequence 7, Application US/09697123B

Initial Score = 137 Optimized Score = 165 Significance = -0.31  
Residue Identity = 80% Matches = 172 Mismatches = 36  
Gaps = 6 Conservative Substitutions = 0

```
X 10 20 30 40 50 60
TCAAGGAGAAGCGCTACGACCTCGCGGGTGGCGCTACAAGGTCAACAAGAGCTCGGCT-----GA
|||||
TCAAGGAGAAGCGCTACGACCTGGCCGGTGGCGGCTGACAGGTGACAAAGAGCTGGTCTTGGCGGTG
X 10 20 30 40 50 60 70

70 80 90 100 110 120 130
ACACCGGTCCTCCGATCACCACGACCACTCTGACCGAAGAGGACGTCGTGCGCCACCATCGAGTACCTGCTCC
|||||
CCAAACCGGCTCTGCTGACCTGACCAACGCTACCGAGGAGAGCGTCGTGCGCCACCATCGGCTACCTGCTGTC
80 90 100 110 120 130 140

140 150 160 170 180 190 200 X
GCCTGCACGAGGCGCACACGATCACCCTCCCGGCGGAGTCGAGGTCCCGGTGGAAACCGACACAT
|||||
GCCTGCACGAGGCGCACACGATGACCCCGCGGCTCGAGGTCCCGGTGCGAGTTCGAGTTCGACACAT
150 160 170 180 190 200 210 X
```

## 21. US-09-697-123B-17 (1-208)

US-09-697-123B-9 Sequence 9, Application US/09697123B

Initial Score = 137 Optimized Score = 164 Significance = -0.31  
Residue Identity = 78% Matches = 174 Mismatches = 34  
Gaps = 15 Conservative Substitutions = 0

```
X 10 20 30 40 50 60 70
TCAAGGAGAAGCGCTACGACCTCGCGGGTGGCGCTACAAGGTCAACAAGAGCTCGGCTGAACACCG
|||||
TCAAGGAGAAGCGCTACGACCTGGCCCGCTCGCCGCTACAAGGTCAACAAGAGCTGGTCTGACGCGCG
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140
CGTCCCGGATCACCACGACCACTGACCGAAGAGGACGTCGTGCGCCACCATCGAGTACTGTCGCGCTGC
|||||
GCGAGCGGATCAGCTGTCACGCTGACCGGAGGAGACGTCGTGCGCCACCATCGAATACCTGGTCTGCGCTGC
80 90 100 110 120 130 140

150 160 170 180 190 200
ACGAGGCG-----CACACCACGATGACCGTCCCGGGGAGTCGAGTTCGCGGTGGAAACCG
|||||
ACCACGCGGCTACGAGTGGCGAGCGCGCGTATGACTGTCCCGCGGGCATCGAGTTCGCGGTGGAGACCG
150 160 170 180 190 200 210
```

```
X
ACGACAT
|||||
ACGACAT
220 X
```

## 22. US-09-697-123B-17 (1-208)

US-09-697-123B-1 Sequence 18, Application US/09697123B

Initial Score = 113 Optimized Score = 169 Significance = -0.80  
Residue Identity = 82% Matches = 175 Mismatches = 33  
Gaps = 3 Conservative Substitutions = 0

```
X 10 20 30 40 50 60
TCAAGGAGAAGCGCTACGACCTCGCGGGTGGCGCTACAAGGTCAACAAGAGCTCGGCTGA-----ACA
|||||
TCAAGGAGAAGCGCTACGACCTGGCCCGCTGGGTGCGGTACAAGGTCAACAAGAGCTGGGCTACCGGAGA
X 10 20 30 40 50 60 70

70 80 90 100 110 120 130 140
CCGCTCCCGGATCAGACGACCACTTGACCGAAGAGGAGCTGTCGCCACCATCGAGTACTGTCGTCGCGC
|||||
ACCCGCGCGACACGACCTCGACGCTGACCGTGAACGAGAGGAGCTGTCGCCACCATCGAGTACTGTCGCGC
80 90 100 110 120 130 140

150 160 170 180 190 200 X
TCCAGAGGGGCGCACACGACGATGACCGTCCCGGGGAGTCGAGGTCCGGTGGAAACCGACGACAT
|||||
TGCATCAGGGGCGCACAGACGATGACCGTCCCGGGTGGAGTCGAGTCCCGTTCGAGGTTCGACGACAT
150 160 170 180 190 200 210
```

## 23. US-09-697-123B-17 (1-208)

US-09-697-123B-2 Sequence 21, Application US/09697123B

Initial Score = 111 Optimized Score = 169 Significance = -0.84  
Residue Identity = 82% Matches = 176 Mismatches = 32  
Gaps = 6 Conservative Substitutions = 0

```
X 10 20 30 40 50 60
TCAAGGAGAAGCGCTACGACCTCGCGGGTGGCGCTACAAGGTCAACAAGAGCTCGGCTG-----A
|||||
TCAAGGAGAAGCGCTACGATCTGCGCCGCTGGTTCGGTACAAGGTGAACAAGAGCTGGCGTGGCGGCA
X 10 20 30 40 50 60 70

70 80 90 100 110 120 130
ACACCGGTCCTCCGATCAGACGACCACTTGACCGAAGAGGACGTCGTGCGCACCATCGAGTACTGTCCTC
|||||
CCAAATCCGGCTCAGGTGACCAACGACCACTTCAACGAGGAGACGTCGTGCGCACCATCGAGTACTGTCCTC
80 90 100 110 120 130 140

140 150 160 170 180 190 200 X
GCCTGCACGAGGCGCACACGATGACCGTCCCGGGGAGTCGAGGTGCGGTGGAAACCGACGACAT
|||||
GCCTGCACGAGGCGCACACGATGACCGTCCCGGGCGGCTCGAGGTGCGGTGTCGAGTTCGAGTTCGAGTTCG
150 160 170 180 190 200 210 X
```

## 24. US-09-697-123B-17 (1-208)

US-09-697-123B-1 Sequence 14, Application US/09697123B

Initial Score = 109 Optimized Score = 169 Significance = -0.88  
Residue Identity = 82% Matches = 176 Mismatches = 32  
Gaps = 6 Conservative Substitutions = 0

```
X 10 20 30 40 50 60 70
TCAAGGAGAAGCGCTACGACCTCGCGGGTGGCGCTACAAGGTCAACAAGAGCTCGGCTGAACACCG
|||||
TCAAGGAGAAGCGCTACGACCTGGCCCGGTTGGCGGTTCACAGGTTCACAAAGAGCTCGGCTGCCGCGCG
X 10 20 30 40 50 60 70
```





X 10 20 30 40 50 60 70  
GCAGGCCATCAGCTGCGACGCTGACCGAAGAAAGAGCTGCGCCACCATCGAATATCTGTCGCTTGC  
GCAGGCCATCAGCTGCGACGCTGACCGAAGAAAGAGCTGCGCCACCATCGAATATCTGTCGCTTGC  
GCAGGCCATCAGCTGCGACGCTGACCGAAGAAAGAGCTGCGCCACCATCGAATATCTGTCGCTTGC  
150 160 170 180 190 200 X  
ACGAGGCTCAGACACCATGATGACCGCTGCGCGCGCGCTGCGAGGTGCGCGGGAACCGACGACAT  
ACGAGGCTCAGACACCATGATGACCGCTGCGCGCGCGCTGCGAGGTGCGCGGGAACCGACGACAT  
ACGAGGCTCAGACACCATGATGACCGCTGCGCGCGCGCTGCGAGGTGCGCGGGAACCGACGACAT  
150 160 170 180 190 200 X

3. US-09-697-123B-16 (1-208)  
US-09-697-123B-5 Sequence 5, Application US/09697123B

Initial Score = 208 Optimized Score = 208 Significance = 1.02  
Residue Identity = 100% Matches = 208 Mismatches = 0  
Gaps = 0 Conservative Substitutions = 0  
X 10 20 30 40 50 60 70  
TCAAGAGAAAGCGCTAGACGCTGCGCGCGCGCTGCGCTATTAAGTCAACAAGAAAGCTGCGCTGCTGCG  
TCAAGAGAAAGCGCTAGACGCTGCGCGCGCGCTGCGCTATTAAGTCAACAAGAAAGCTGCGCTGCTGCG  
X 10 20 30 40 50 60 70  
GCAGGCCATCAGCTGCGACGCTGACCGAAGAAAGAGCTGCGCCACCATCGAATATCTGTCGCTTGC  
GCAGGCCATCAGCTGCGACGCTGACCGAAGAAAGAGCTGCGCCACCATCGAATATCTGTCGCTTGC  
GCAGGCCATCAGCTGCGACGCTGACCGAAGAAAGAGCTGCGCCACCATCGAATATCTGTCGCTTGC  
80 90 100 110 120 130 140  
ACGAGGCTCAGACACCATGATGACCGCTGCGCGCGCGCTGCGAGGTGCGCGGGAACCGACGACAT  
ACGAGGCTCAGACACCATGATGACCGCTGCGCGCGCGCTGCGAGGTGCGCGGGAACCGACGACAT  
ACGAGGCTCAGACACCATGATGACCGCTGCGCGCGCGCTGCGAGGTGCGCGGGAACCGACGACAT  
150 160 170 180 190 200 X

4. US-09-697-123B-16 (1-208)  
US-09-697-123B-2 Sequence 22, Application US/09697123B

Initial Score = 207 Optimized Score = 207 Significance = 1.00  
Residue Identity = 99% Matches = 207 Mismatches = 1  
Gaps = 0 Conservative Substitutions = 0  
X 10 20 30 40 50 60 70  
TCAAGAGAAAGCGCTAGACGCTGCGCGCGCGCTGCGCTATTAAGTCAACAAGAAAGCTGCGCTGCTGCG  
TCAAGAGAAAGCGCTAGACGCTGCGCGCGCGCTGCGCTATTAAGTCAACAAGAAAGCTGCGCTGCTGCG  
X 10 20 30 40 50 60 70  
GCAGGCCATCAGCTGCGACGCTGACCGAAGAAAGAGCTGCGCCACCATCGAATATCTGTCGCTTGC  
GCAGGCCATCAGCTGCGACGCTGACCGAAGAAAGAGCTGCGCCACCATCGAATATCTGTCGCTTGC  
GCAGGCCATCAGCTGCGACGCTGACCGAAGAAAGAGCTGCGCCACCATCGAATATCTGTCGCTTGC  
80 90 100 110 120 130 140  
ACGAGGCTCAGACACCATGATGACCGCTGCGCGCGCGCTGCGAGGTGCGCGGGAACCGACGACAT  
ACGAGGCTCAGACACCATGATGACCGCTGCGCGCGCGCTGCGAGGTGCGCGGGAACCGACGACAT  
ACGAGGCTCAGACACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT  
150 160 170 180 190 200 X

5. US-09-697-123B-16 (1-208)  
US-09-697-123B-1 Sequence 10, Application US/09697123B

Initial Score = 186 Optimized Score = 186 Significance = 0.60  
Residue Identity = 89% Matches = 186 Mismatches = 22  
Gaps = 0 Conservative Substitutions = 0

X 10 20 30 40 50 60 70  
TCAAGAGAAAGCGCTAGACGCTGCGCGCGCGCTGCGCTATTAAGTCAACAAGAAAGCTGCGCTGCTGCG  
TCAAGAGAAAGCGCTAGACGCTGCGCGCGCGCTGCGCTATTAAGTCAACAAGAAAGCTGCGCTGCTGCG  
X 10 20 30 40 50 60 70  
GCAGGCCATCAGCTGCGACGCTGACCGAAGAAAGAGCTGCGCCACCATCGAATATCTGTCGCTTGC  
GCAGGCCATCAGCTGCGACGCTGACCGAAGAAAGAGCTGCGCCACCATCGAATATCTGTCGCTTGC  
GCAGGCCATCAGCTGCGACGCTGACCGAAGAAAGAGCTGCGCCACCATCGAATATCTGTCGCTTGC  
80 90 100 110 120 130 140  
ACGAGGCTCAGACACCATGATGACCGCTGCGCGCGCGCTGCGAGGTGCGCGGGAACCGACGACAT  
ACGAGGCTCAGACACCATGATGACCGCTGCGCGCGCGCTGCGAGGTGCGCGGGAACCGACGACAT  
ACGAGGCTCAGACACCATGATGACCGCTGCGCGCGCGCTGCGAGGTGCGCGGGAACCGACGACAT  
150 160 170 180 190 200 X

6. US-09-697-123B-16 (1-208)  
US-09-697-123B-1 Sequence 11, Application US/09697123B

Initial Score = 185 Optimized Score = 185 Significance = 0.58  
Residue Identity = 88% Matches = 185 Mismatches = 23  
Gaps = 0 Conservative Substitutions = 0  
X 10 20 30 40 50 60 70  
TCAAGAGAAAGCGCTAGACGCTGCGCGCGCGCTGCGCTATTAAGTCAACAAGAAAGCTGCGCTGCTGCG  
TCAAGAGAAAGCGCTAGACGCTGCGCGCGCGCTGCGCTATTAAGTCAACAAGAAAGCTGCGCTGCTGCG  
X 10 20 30 40 50 60 70  
GCAGGCCATCAGCTGCGACGCTGACCGAAGAAAGAGCTGCGCCACCATCGAATATCTGTCGCTTGC  
GCAGGCCATCAGCTGCGACGCTGACCGAAGAAAGAGCTGCGCCACCATCGAATATCTGTCGCTTGC  
GCAGGCCATCAGCTGCGACGCTGACCGAAGAAAGAGCTGCGCCACCATCGAATATCTGTCGCTTGC  
80 90 100 110 120 130 140  
ACGAGGCTCAGACACCATGATGACCGCTGCGCGCGCGCTGCGAGGTGCGCGGGAACCGACGACAT  
ACGAGGCTCAGACACCATGATGACCGCTGCGCGCGCGCTGCGAGGTGCGCGGGAACCGACGACAT  
ACGAGGCTCAGACACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT  
150 160 170 180 190 200 X

7. US-09-697-123B-16 (1-208)  
US-09-697-123B-3 Sequence 3, Application US/09697123B

Initial Score = 185 Optimized Score = 185 Significance = 0.58  
Residue Identity = 88% Matches = 185 Mismatches = 23  
Gaps = 0 Conservative Substitutions = 0  
X 10 20 30 40 50 60 70  
TCAAGAGAAAGCGCTAGACGCTGCGCGCGCGCTGCGCTATTAAGTCAACAAGAAAGCTGCGCTGCTGCG  
TCAAGAGAAAGCGCTAGACGCTGCGCGCGCGCTGCGCTATTAAGTCAACAAGAAAGCTGCGCTGCTGCG  
X 10 20 30 40 50 60 70  
GCAGGCCATCAGCTGCGACGCTGACCGAAGAAAGAGCTGCGCCACCATCGAATATCTGTCGCTTGC  
GCAGGCCATCAGCTGCGACGCTGACCGAAGAAAGAGCTGCGCCACCATCGAATATCTGTCGCTTGC  
GCAGGCCATCAGCTGCGACGCTGACCGAAGAAAGAGCTGCGCCACCATCGAATATCTGTCGCTTGC  
80 90 100 110 120 130 140  
ACGAGGCTCAGACACCATGATGACCGCTGCGCGCGCGCTGCGAGGTGCGCGGGAACCGACGACAT  
ACGAGGCTCAGACACCATGATGACCGCTGCGCGCGCGCTGCGAGGTGCGCGGGAACCGACGACAT  
ACGAGGCTCAGACACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT  
150 160 170 180 190 200 X

8. US-09-697-123B-16 (1-208)



## US-09-697-123B-1 Sequence 15, Application US/09697123B

Initial Score = 184 Optimized Score = 184 Significance = 0.56  
Residue Identity = 88% Matches = 184 Mismatches = 24  
Gaps = 0 Conservative Substitutions = 0

X  
TCGAAGAGAAAGCCCTACGACGCTGCGCCGCGGCGGCTGCTATAGGTCAACAAAGAGCTCGGGCTGCATCTGC  
|||||  
TCGAAGAGAAAGCCCTACGACGCTGCGCCGCGGCGGCTGCTATAGGTCAACAAAGAGCTCGGGCTGCATCTGC  
X  
10 20 30 40 50 60 70

80 90 100 110 120 130 140  
GCGAGCCCATCAGCTGCTGACGCTGACGCGGCGGCGGCTGCTATAGGTCAACAAAGAGCTCGGGCTGCATCTGC  
|||||  
GTCAGGCGATCAGACGCTGACGCTGACGCGGCGGCGGCTGCTATAGGTCAACAAAGAGCTCGGGCTGCATCTGC  
X  
10 20 30 40 50 60 70

150 160 170 180 190 200  
ACGAGGCTCAGACGACGATGACGCTGCTGCGGCGGCGGCTGCTATAGGTCAACAAAGAGCTCGGGCTGCATCTGC  
|||||  
ACGAGGCTCAGACGACGATGACGCTGCTGCGGCGGCGGCTGCTATAGGTCAACAAAGAGCTCGGGCTGCATCTGC  
X  
150 160 170 180 190 200

## 9. US-09-697-123B-16 (1-208)

US-09-697-123B-2 Sequence 23, Application US/09697123B

Initial Score = 182 Optimized Score = 182 Significance = 0.52  
Residue Identity = 87% Matches = 182 Mismatches = 26  
Gaps = 0 Conservative Substitutions = 0

X  
TCGAAGAGAAAGCCCTACGACGCTGCGCCGCGGCGGCTGCTATAGGTCAACAAAGAGCTCGGGCTGCATCTGC  
|||||  
TCGAAGAGAAAGCCCTACGACGCTGCGCCGCGGCGGCTGCTATAGGTCAACAAAGAGCTCGGGCTGCATCTGC  
X  
10 20 30 40 50 60 70

80 90 100 110 120 130 140  
GCGAGCCCATCAGCTGCTGACGCTGACGCGGCGGCGGCTGCTATAGGTCAACAAAGAGCTCGGGCTGCATCTGC  
|||||  
GTCAGGCGATCAGACGCTGACGCTGACGCGGCGGCGGCTGCTATAGGTCAACAAAGAGCTCGGGCTGCATCTGC  
X  
80 90 100 110 120 130 140

150 160 170 180 190 200  
ACGAGGCTCAGACGACGATGACGCTGCTGCGGCGGCGGCTGCTATAGGTCAACAAAGAGCTCGGGCTGCATCTGC  
|||||  
ATGAGGCTCAGACGACGATGACGCTGCTGCGGCGGCGGCTGCTATAGGTCAACAAAGAGCTCGGGCTGCATCTGC  
X  
150 160 170 180 190 200

## 10. US-09-697-123B-16 (1-208)

US-09-697-123B-1 Sequence 19, Application US/09697123B

Initial Score = 182 Optimized Score = 182 Significance = 0.52  
Residue Identity = 87% Matches = 182 Mismatches = 26  
Gaps = 0 Conservative Substitutions = 0

X  
TCGAAGAGAAAGCCCTACGACGCTGCGCCGCGGCGGCTGCTATAGGTCAACAAAGAGCTCGGGCTGCATCTGC  
|||||  
TCGAAGAGAAAGCCCTACGACGCTGCGCCGCGGCGGCTGCTATAGGTCAACAAAGAGCTCGGGCTGCATCTGC  
X  
10 20 30 40 50 60 70

80 90 100 110 120 130 140  
GCGAGCCCATCAGCTGCTGACGCTGACGCGGCGGCGGCTGCTATAGGTCAACAAAGAGCTCGGGCTGCATCTGC  
|||||  
GTCAGGCGATCAGACGCTGACGCTGACGCGGCGGCGGCTGCTATAGGTCAACAAAGAGCTCGGGCTGCATCTGC  
X  
80 90 100 110 120 130 140

150 160 170 180 190 200  
ACGAGGCTCAGACGACGATGACGCTGCTGCGGCGGCGGCTGCTATAGGTCAACAAAGAGCTCGGGCTGCATCTGC  
|||||  
ACGAGGCTCAGACGACGATGACGCTGCTGCGGCGGCGGCTGCTATAGGTCAACAAAGAGCTCGGGCTGCATCTGC  
X  
150 160 170 180 190 200

ACGAGGCTCAGACGACGATGACGCTGCTGCGGCGGCGGCTGCTATAGGTCAACAAAGAGCTCGGGCTGCATCTGC  
X  
150 160 170 180 190 200

## 11. US-09-697-123B-8 Sequence 8, Application US/09697123B

Initial Score = 182 Optimized Score = 182 Significance = 0.52  
Residue Identity = 87% Matches = 182 Mismatches = 26  
Gaps = 0 Conservative Substitutions = 0

X  
TCGAAGAGAAAGCCCTACGACGCTGCGCCGCGGCGGCTGCTATAGGTCAACAAAGAGCTCGGGCTGCATCTGC  
|||||  
TCGAAGAGAAAGCCCTACGACGCTGCGCCGCGGCGGCTGCTATAGGTCAACAAAGAGCTCGGGCTGCATCTGC  
X  
10 20 30 40 50 60 70

80 90 100 110 120 130 140  
GCGAGCCCATCAGCTGCTGACGCTGACGCGGCGGCGGCTGCTATAGGTCAACAAAGAGCTCGGGCTGCATCTGC  
|||||  
ATCATCCGATCAGACGACGCTGACGCGGCGGCGGCTGCTATAGGTCAACAAAGAGCTCGGGCTGCATCTGC  
X  
80 90 100 110 120 130 140

150 160 170 180 190 200  
ACGAGGCTCAGACGACGATGACGCTGCTGCGGCGGCGGCTGCTATAGGTCAACAAAGAGCTCGGGCTGCATCTGC  
|||||  
ACGAGGCTCAGACGACGATGACGCTGCTGCGGCGGCGGCTGCTATAGGTCAACAAAGAGCTCGGGCTGCATCTGC  
X  
150 160 170 180 190 200

## 12. US-09-697-123B-16 (1-208)

US-09-697-123B-1 Sequence 1, Application US/09697123B

Initial Score = 182 Optimized Score = 182 Significance = 0.52  
Residue Identity = 87% Matches = 182 Mismatches = 26  
Gaps = 0 Conservative Substitutions = 0

X  
TCGAAGAGAAAGCCCTACGACGCTGCGCCGCGGCGGCTGCTATAGGTCAACAAAGAGCTCGGGCTGCATCTGC  
|||||  
TCGAAGAGAAAGCCCTACGACGCTGCGCCGCGGCGGCTGCTATAGGTCAACAAAGAGCTCGGGCTGCATCTGC  
X  
10 20 30 40 50 60 70

80 90 100 110 120 130 140  
GCGAGCCCATCAGCTGCTGACGCTGACGCGGCGGCGGCTGCTATAGGTCAACAAAGAGCTCGGGCTGCATCTGC  
|||||  
GTCAGGCGATCAGACGCTGACGCTGACGCGGCGGCGGCTGCTATAGGTCAACAAAGAGCTCGGGCTGCATCTGC  
X  
80 90 100 110 120 130 140

150 160 170 180 190 200  
ACGAGGCTCAGACGACGATGACGCTGCTGCGGCGGCGGCTGCTATAGGTCAACAAAGAGCTCGGGCTGCATCTGC  
|||||  
ACGAGGCTCAGACGACGATGACGCTGCTGCGGCGGCGGCTGCTATAGGTCAACAAAGAGCTCGGGCTGCATCTGC  
X  
150 160 170 180 190 200

## 13. US-09-697-123B-16 (1-208)

US-09-697-123B-2 Sequence 2, Application US/09697123B

Initial Score = 181 Optimized Score = 181 Significance = 0.50  
Residue Identity = 87% Matches = 181 Mismatches = 27  
Gaps = 0 Conservative Substitutions = 0

X  
TCGAAGAGAAAGCCCTACGACGCTGCGCCGCGGCGGCTGCTATAGGTCAACAAAGAGCTCGGGCTGCATCTGC  
|||||  
TCGAAGAGAAAGCCCTACGACGCTGCGCCGCGGCGGCTGCTATAGGTCAACAAAGAGCTCGGGCTGCATCTGC  
X  
10 20 30 40 50 60 70

80 90 100 110 120 130 140  
GCGAGCCCATCAGCTGCTGACGCTGACGCGGCGGCGGCTGCTATAGGTCAACAAAGAGCTCGGGCTGCATCTGC  
|||||  
GTCAGGCGATCAGACGCTGACGCTGACGCGGCGGCGGCTGCTATAGGTCAACAAAGAGCTCGGGCTGCATCTGC  
X  
80 90 100 110 120 130 140

[illegible]

14. US-09-697-123B-16 (1-208)	US-09-697-123B-1 Sequence 17, Application US/09697123B				
Initial Score =	178	Optimized Score =	178	Significance =	0.44
Residue Identity =	85%	Matches	178	Mismatches	30
Gaps	0	Conservative Substitutions	-	-	0

X 10 20 30 40 50 60 70  
 TCAAGAGGAAGCGCTACGACCGGCCCGCGTGGCTATTAAGTCAACGAAGTGGGCTGCATGTGCG  
 X 10 20 30 40 50 60 70  
 TCAAGAGGAAGCGCTACGACCTCGCGGGGTGGCCGTACAAAGTCAACGAAGTGGGCTGCATGCACCG  
 X 10 20 30 40 50 60 70  
 GCGAGCCCATCAGCTGCTGCGACGCTGCCCAAGAAACCTGGCCGCCACCTGCAATATCTGTCCTGTC  
 80 90 100 110 120 130 140  
 CGTCCCGCATCAGACGACCACTGTGACCCCAAGAGACCTTCCGCCACCTGCAAGTACTGGTCCGCTTGC  
 80 90 100 110 120 130 140  
 150 160 170 180 190 200 X  
 ACGAGGCTGAGCACGATGACCTTCCGGGCGCGCTCAGGTGGCGTGAACAGCAGACAT  
 150 160 170 180 190 200 X  
 ACGAGGCGCACACGATGACGCTCCGGGCGGAGATCGAGTGGCGGTGAACCGCAGACAT  
 150 160 170 180 190 200 X

Initial Score	Residue Identity	Optimized Score	Matches	Mismatches	Significance
-	82%	-	172	172	0.33
	=	0	Conservative Substitutions	=	36
					0

US-09-697-123B-16 (1-208)  
US-09-697-123B-2 Sequence 24, Application US/09697123B

X	10	20	30	40	50	60	70
TCACGAGCAAGCGCTACGACCTGGCGCCGCTGCTTAAAGTCACAAAGCAAGCAAGCGGCTGCATATG							
TCACGAGCAAGCGCTACGACCTGGCGCCGCTGCTTAAAGTCACAAAGCAAGCAAGCGGCTGCATATG							
X	10	20	30	40	50	60	70

GCACGCCCATCTCATCTGCTCCAGCGTACAGCGAAGAACGCTGTCGGCCACCATGAAATCTTGCTGGCTTGC  
 150 160 170 180 190 200 210 220 230 240  
 AGAATCTGGCGAACACACACACCGCTGACCGGAGAGACAGCTGTCGCCACCATCGAATTAAGTCTGTCGGCTGC  
 80 90 100 110 120 130 140  
 150 160 170 180 190 200 210 220 230 240  
 ACCGAGGCTCAGACGACGATGACCGTTCGGGGGGGCTGAGAGTCCCGCTGGAAGCAACCGAGACAT  
 150 160 170 180 190 200 210 220 230 240  
 ACAGGGGACGCCCAACGATGAAGTCCCGGCTGGAGTCCGAGTGGAGCAACCGAGACAT X  
 150 160 170 180 190 200 210 220 230 240

```

16. US-09-697-123B-16 (1-208)
US-09-697-123B-2 Sequence 20, Application US/09697123B

Initial Score = 161 Optimized Score = 171 Significance = 0.12
Residue Identity = 85% Matches = 177 Mismatches = 28
Gaps = 3 Conservative Substitutions = 0

X      10      20      30      40      50      60      70
TCAAGCAAGACGGCTAGACACTGGCCGGCGCTGCGCATATAGTCAACAAGAAGCTGGCGCTGCATGTCG
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
TCAAGCAAGACCGCTACGACACTGGCGGTGTCGGCCGCTCAACAGGTCAACAAGAAGCTGCACGCGGG
X      10      20      30      40      50      60      70

```

[illegible][illegible]

Initial Score	-	157	Optimized Score	-	176	Significance	-	0.04
Residue Identity	-	87%	Matches	-	181	Mismatches	-	26
Gaps	-	1	Conservative Substitutions	-	0		-	0

[illegible][illegible]

```

18. US-09-697-123B-16 (1-208)
   US-09-697-123B-4 Sequence 4, Application US/09657123B

Initial Score      = 147      Optimized Score      = 178      Significance      = -0.15
Residue Identity   = 87%      Matches          = 183      Mismatches        = 24
Gaps               = 1        Conservative Substitutions = 0

```

[illegible][illegible]

	150	160	170	180	190	200	
ACGAGAGGTCAGACCCAGTACCGTTCGGGGCGCGTCGAGTGGCGGTAACGAGACAT							x
ACGAGAGGTCAGACCCAGTACCGTTCGGGGCGCGTCGAGTGGCGGTAACGAGACAT	150	160	170	180	190	200	
ACGAGAGGTCAGACCCAGTACCGTTCGGGGCGCGTCGAGTGGCGGTAACGAGACAT	150	160	170	180	190	200	x

Initial Score	=	144	Optimized Score	=	170	Significance	=	-0.21
Residue Identity	=	80%	Matches	=	180	Mismatches	=	28
Gaps	=	15	Conservative Substitutions	=			=	0

X 10 20 30 40 50 60 70  
TCAAGAGAGAGCCCTACGACCTGGCGCCGCTCGTGTAAAGTCACCAAGAGCTCCGGCTGCATATGCG  
|||||  
TCAAGAGAGAGCCCTACGACCTGGCGCCGCTCGTGTAAAGTCACCAAGAGCTCCGGCTGCATATGCG  
|||||  
X 10 20 30 40 50 60 70  
CGGAGGCC-----CATCATGCTGCTGACACCTGACCCCAAGAGAGCTGCTGGCCGACCATCGAATATCTGGTCC

X  
10 20 30 40 50 60 70  
TAAAGGAGAAAGCGCTACGACCTGGCCGCCGCTCGCTTAAAGTCAACAAAGACTGGGGCTGCATATGCTC  
TAAAGGAGAGCGCTACGATCTGGCCGCCGCTCGCTCAAGGTGAACAAAGACTGGGGCTGGGGCGCA

```
      X      10      20      30      40      50      60      70
      |      |      |      |      |      |      |
GGGAGCC-----CATCAGTCGTGACGGCTGACCGAAGAAGAGTCGTGCGCCACCATCGAATATCTGGTCC
      |      |      |      |      |      |      |      |      |      |
CCAATCCGGCTCAGGTGACACACACACACCTTCACCGGAAAGAGTCGTGCGCCACCATCGAATCTGGTCC
      80      90     100     110     120     130     140
140      150     160     170     180     190     200
GCTTCCACGAGGGTCAGACGACGATGACCGTTCGGGGCGGCGTCGAGGTGCCGGTGAACCGACGACAT
      |      |      |      |      |      |      |      |      |      |
GCTTGCACGAGGGCGCCAGACGATGACCGCCCGCGCGCGTCGAGGTGCCGGTGAATGTGAGACGACAT
      150     160     170     180     190     200     210
X
```

## 25. US-09-697-123B-16 (1-208)

US-09-697-123B-2 Sequence 25, Application US/09697123B

```
Initial Score      = 19 Optimized Score = 19 Significance = -2.62
Residue Identity = 100% Matches = 19 Mismatches = 0
Gaps = 0 Conservative Substitutions = 0
```

```
X      10
TCAAGAGAGAAGCGCTACGA
      |      |      |      |      |
TCAAGAGAGAAGCGCTACGA
      X      10
```



8. US-09-697-123B-15 (1-208)  
US-09-697-123B-1 Sequence 10, Application US/09697123B

Initial Score = 188 Optimized Score = 188 Significance = 0.59  
Residue Identity = 90% Matches = 188 Mismatches = 20  
Gaps = 0 Conservative Substitutions = 0

X 10 20 30 40 50 60 70  
TCAAGCAGAACCGCTACGACCTGGCCCGGGTGGCCCGCTACAGGTCAACAAAGCTGACGCTGCACGCGC  
TCAAGCAGAACCGCTACGACCTGGCCCGGGTGGCCCGCTACAGGTCAACAAAGCTGACGCTGCACGCGC  
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140  
GTGAGCCGATCACCAGCTCGACGCTGACCGAGAAAGAGTCTGCGCCACCATCGATCGCTGCTGCGCTGC  
GCCAGCCATCACCAGCTCGACGCTGACCGAGAAAGAGTCTGCGCCACCATCGATCGCTGCTGCGCTGC  
X 80 90 100 110 120 130 140

150 160 170 180 190 200 X  
ACGAGGCTCAGCCCGACGATGACCGTCCCGCGGCGATCGAGGTCCCGGTGAGAGCCGACGACAT  
ACGAGGCTCAGCCCGACGATGACCGTCCCGCGGCGATCGAGGTCCCGGTGAGAGCCGACGACAT  
X 150 160 170 180 190 200

9. US-09-697-123B-15 (1-208)  
US-09-697-123B-8 Sequence 8, Application US/09697123B

Initial Score = 185 Optimized Score = 185 Significance = 0.53  
Residue Identity = 88% Matches = 185 Mismatches = 23  
Gaps = 0 Conservative Substitutions = 0

X 10 20 30 40 50 60 70  
TCAAGCAGAACCGCTACGACCTGGCCCGGGTGGCCCGCTACAGGTCAACAAAGCTGCGCTGCACGCGC  
TCAAGCAGAACCGCTACGACCTGGCCCGGGTGGCCCGCTACAGGTCAACAAAGCTGCGCTGCACGCGC  
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140  
GTGAGCCGATCACCAGCTCGACGCTGACCGAGAAAGAGTCTGCGCCACCATCGATCGCTGCTGCGCTGC  
ATCATCCGATCACCAGCTCGACGCTGACCGAGAAAGAGTCTGCGCCACCATCGATCGCTGCTGCGCTGC  
X 80 90 100 110 120 130 140

150 160 170 180 190 200 X  
ACGAGGCTCAGCCCGACGATGACCGTCCCGCGGCGATCGAGGTCCCGGTGAGAGCCGACGACAT  
ACGAGGCTCAGCCCGACGATGACCGTCCCGCGGCGATCGAGGTCCCGGTGAGAGCCGACGACAT  
X 150 160 170 180 190 200

10. US-09-697-123B-15 (1-208)  
US-09-697-123B-1 Sequence 16, Application US/09697123B

Initial Score = 184 Optimized Score = 184 Significance = 0.51  
Residue Identity = 88% Matches = 184 Mismatches = 24  
Gaps = 0 Conservative Substitutions = 0

X 10 20 30 40 50 60 70  
TCAAGCAGAACCGCTACGACCTGGCCCGGGTGGCCCGCTACAGGTCAACAAAGCTGCGCTGCACGCGC  
TCAAGCAGAACCGCTACGACCTGGCCCGGGTGGCCCGCTACAGGTCAACAAAGCTGCGCTGCACGCGC  
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140  
GTGAGCCGATCACCAGCTCGACGCTGACCGAGAAAGAGTCTGCGCCACCATCGATCGCTGCTGCGCTGC  
GCCAGCCATCACCAGCTCGACGCTGACCGAGAAAGAGTCTGCGCCACCATCGATCGCTGCTGCGCTGC  
X 80 90 100 110 120 130 140

150 160 170 180 190 200 X  
ACGAGGCTCAGCCCGACGATGACCGTCCCGCGGCGATCGAGGTCCCGGTGAGAGCCGACGACAT  
ACGAGGCTCAGCCCGACGATGACCGTCCCGCGGCGATCGAGGTCCCGGTGAGAGCCGACGACAT  
X 150 160 170 180 190 200

11. US-09-697-123B-15 (1-208)  
US-09-697-123B-6 Sequence 6, Application US/09697123B

Initial Score = 184 Optimized Score = 184 Significance = 0.51  
Residue Identity = 88% Matches = 184 Mismatches = 24  
Gaps = 0 Conservative Substitutions = 0

X 10 20 30 40 50 60 70  
TCAAGCAGAACCGCTACGACCTGGCCCGGGTGGCCCGCTACAGGTCAACAAAGCTGCGCTGCACGCGC  
TCAAGCAGAACCGCTACGACCTGGCCCGGGTGGCCCGCTACAGGTCAACAAAGCTGCGCTGCACGCGC  
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140  
GTGAGCCGATCACCAGCTCGACGCTGACCGAGAAAGAGTCTGCGCCACCATCGATCGCTGCTGCGCTGC  
GCCAGCCATCACCAGCTCGACGCTGACCGAGAAAGAGTCTGCGCCACCATCGATCGCTGCTGCGCTGC  
X 80 90 100 110 120 130 140

150 160 170 180 190 200 X  
ACGAGGCTCAGCCCGACGATGACCGTCCCGCGGCGATCGAGGTCCCGGTGAGAGCCGACGACAT  
ACGAGGCTCAGCCCGACGATGACCGTCCCGCGGCGATCGAGGTCCCGGTGAGAGCCGACGACAT  
X 150 160 170 180 190 200

12. US-09-697-123B-15 (1-208)  
US-09-697-123B-5 Sequence 5, Application US/09697123B

Initial Score = 184 Optimized Score = 184 Significance = 0.51  
Residue Identity = 88% Matches = 184 Mismatches = 24  
Gaps = 0 Conservative Substitutions = 0

X 10 20 30 40 50 60 70  
TCAAGCAGAACCGCTACGACCTGGCCCGGGTGGCCCGCTACAGGTCAACAAAGCTGCGCTGCACGCGC  
TCAAGCAGAACCGCTACGACCTGGCCCGGGTGGCCCGCTACAGGTCAACAAAGCTGCGCTGCACGCGC  
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140  
GTGAGCCGATCACCAGCTCGACGCTGACCGAGAAAGAGTCTGCGCCACCATCGATCGCTGCTGCGCTGC  
GGGAGCCATCACCAGCTCGACGCTGACCGAGAAAGAGTCTGCGCCACCATCGATCGCTGCTGCGCTGC  
X 80 90 100 110 120 130 140

150 160 170 180 190 200 X  
ACGAGGCTCAGCCCGACGATGACCGTCCCGCGGCGATCGAGGTCCCGGTGAGAGCCGACGACAT  
ACGAGGCTCAGCCCGACGATGACCGTCCCGCGGCGATCGAGGTCCCGGTGAGAGCCGACGACAT  
X 150 160 170 180 190 200

13. US-09-697-123B-15 (1-208)  
US-09-697-123B-2 Sequence 22, Application US/09697123B

Initial Score = 183 Optimized Score = 183 Significance = 0.49  
Residue Identity = 87% Matches = 183 Mismatches = 25  
Gaps = 0 Conservative Substitutions = 0

X 10 20 30 40 50 60 70  
TCAAGCAGAACCGCTACGACCTGGCCCGGGTGGCCCGCTACAGGTCAACAAAGCTGCGCTGCACGCGC  
TCAAGCAGAACCGCTACGACCTGGCCCGGGTGGCCCGCTACAGGTCAACAAAGCTGCGCTGCACGCGC  
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140  
GTGAGCCGATCACCAGCTCGACGCTGACCGAGAAAGAGTCTGCGCCACCATCGATCGCTGCTGCGCTGC  
GCCAGCCATCACCAGCTCGACGCTGACCGAGAAAGAGTCTGCGCCACCATCGATCGCTGCTGCGCTGC  
X 80 90 100 110 120 130 140

150 160 170 180 190 200 X  
ACGAGGCTACACCCACGATGACCGTCCCCCGGCGATCGAGTCCCGGTGAGAGCGACGACAT  
|||||  
ACGAGGCTACACCCACGATGATGTTCCGGCGGCGATCGAGTCCCGGTGAGAGCGACGACAT  
150 160 170 180 190 200 X

14. US-09-697-123B-15 (1-208)  
US-09-697-123B-1 Sequence 17, Application US/09697123B

Initial Score = 183 Optimized Score = 183 Significance = 0.49  
Residue Identity = 87% Matches = 183 Mismatches = 25  
Gaps = 0 Conservative Substitutions = 0

X 10 20 30 40 50 60 70  
TCAAGAGAAAGCGCTACGACCTGGCCGGGTGGCGGCTCAAGGTCAACAGAAAGCTGCGCTGCACGCG  
|||||  
TCAAGAGAAAGCGCTACGACCTGGCCGGGTGGCGGCTCAAGGTCAACAGAAAGCTGCGCTGCACGCG  
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140  
GTGAGCCGATACACGCTCGACGCTGACCCGAGAGAGCTGCTGCGCACCATTGATCTGCTGCGCTGC  
|||||  
GTGAGCCGATACACGCTCGACGCTGACCCGAGAGAGCTGCTGCGCACCATTGATCTGCTGCGCTGC  
80 90 100 110 120 130 140

150 160 170 180 190 200 X  
ACGAGGCTACACCCACGATGACCGTCCCCCGGCGATCGAGTCCCGGTGAGAGCGACGACAT  
|||||  
ACGAGGCTACACCCACGATGACCGTCCCCCGGCGAGTCAAGGTCAACAGAAAGCTGCGCTGCACGCG  
150 160 170 180 190 200 X

15. US-09-697-123B-15 (1-208)  
US-09-697-123B-2 Sequence 24, Application US/09697123B

Initial Score = 182 Optimized Score = 182 Significance = 0.47  
Residue Identity = 87% Matches = 182 Mismatches = 26  
Gaps = 0 Conservative Substitutions = 0

X 10 20 30 40 50 60 70  
TCAAGAGAAAGCGCTACGACCTGGCCGGGTGGCGGCTCAAGGTCAACAGAAAGCTGCGCTGCACGCG  
|||||  
TCAAGAGAAAGCGCTACGACCTGGCCGGGTGGCGGCTCAAGGTCAACAGAAAGCTGCGCTGCACGCG  
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140  
GTGAGCCGATACACGCTCGACGCTGACCCGAGAGAGCTGCTGCGCACCATTGATCTGCTGCGCTGC  
|||||  
GTGAGCCGATACACGCTCGACGCTGACCCGAGAGAGCTGCTGCGCACCATTGATCTGCTGCGCTGC  
80 90 100 110 120 130 140

150 160 170 180 190 200 X  
ACGAGGCTACACCCACGATGACCGTCCCCCGGCGATCGAGTCCCGGTGAGAGCGACGACAT  
|||||  
ACGAGGCTACACCCACGATGATGTTCCGGCGGCGATCGAGTCCCGGTGAGAGCGACGACAT  
150 160 170 180 190 200 X

16. US-09-697-123B-15 (1-208)  
US-09-697-123B-2 Sequence 20, Application US/09697123B

Initial Score = 177 Optimized Score = 189 Significance = 0.37  
Residue Identity = 93% Matches = 195 Mismatches = 10  
Gaps = 3 Conservative Substitutions = 0

X 10 20 30 40 50 60 70  
TCAAGAGAAAGCGCTACGACCTGGCCGGGTGGCGGCTCAAGGTCAACAGAAAGCTGCGCTGCACGCG  
|||||  
TCAAGAGAAAGCGCTACGACCTGGCCGGGTGGCGGCTCAAGGTCAACAGAAAGCTGCGCTGCACGCG  
X 10 20 30 40 50 60 70

GTGAGCCGATACACGCTCGACGCTGACCCGAGAGAGCTGCTGCGCACCATTGATCTGCTGCGCTGC  
|||||  
GTGAGCCGATACACGCTCGACGCTGACCCGAGAGAGCTGCTGCGCACCATTGATCTGCTGCGCTGC  
80 90 100 110 120 130 140

150 160 170 180 190 200 X  
ACGAGGCTACACCCACGATGACCGTCCCCCGGCGATCGAGTCCCGGTGAGAGCGACGACAT  
|||||  
ACGAGGCTACACCCACGATGATGTTCCGGCGGCGATCGAGTCCCGGTGAGAGCGACGACAT  
150 160 170 180 190 200 X

17. US-09-697-123B-15 (1-208)  
US-09-697-123B-1 Sequence 12, Application US/09697123B

Initial Score = 160 Optimized Score = 183 Significance = 0.04  
Residue Identity = 90% Matches = 188 Mismatches = 19  
Gaps = 1 Conservative Substitutions = 0

X 10 20 30 40 50 60 70  
TCAAGAGAAAGCGCTACGACCTGGCCGGGTGGCGGCTCAAGGTCAACAGAAAGCTGCGCTGCACGCG  
|||||  
TCAAGAGAAAGCGCTACGACCTGGCCGGGTGGCGGCTCAAGGTCAACAGAAAGCTGCGCTGCACGCG  
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140  
GTGAGCCGATACACGCTCGACGCTGACCCGAGAGAGCTGCTGCGCACCATTGATCTGCTGCGCTGC  
|||||  
GTGAGCCGATACACGCTCGACGCTGACCCGAGAGAGCTGCTGCGCACCATTGATCTGCTGCGCTGC  
80 90 100 110 120 130 140

150 160 170 180 190 200 X  
ACGAGGCTACACCCACGATGACCGTCCCCCGGCGATCGAGTCCCGGTGAGAGCGACGACAT  
|||||  
ACGAGGCTACACCCACGATGACCGTCCCCCGGCGAGTCAAGGTCAACAGAAAGCTGCGCTGCACGCG  
150 160 170 180 190 200 X

18. US-09-697-123B-15 (1-208)  
US-09-697-123B-4 Sequence 4, Application US/09697123B

Initial Score = 156 Optimized Score = 185 Significance = -0.04  
Residue Identity = 91% Matches = 190 Mismatches = 17  
Gaps = 1 Conservative Substitutions = 0

X 10 20 30 40 50 60 70  
TCAAGAGAAAGCGCTACGACCTGGCCGGGTGGCGGCTCAAGGTCAACAGAAAGCTGCGCTGCACGCG  
|||||  
TCAAGAGAAAGCGCTACGACCTGGCCGGGTGGCGGCTCAAGGTCAACAGAAAGCTGCGCTGCACGCG  
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140  
GTGAGCCGATACACGCTCGACGCTGACCCGAGAGAGCTGCTGCGCACCATTGATCTGCTGCGCTGC  
|||||  
GTGAGCCGATACACGCTCGACGCTGACCCGAGAGAGCTGCTGCGCACCATTGATCTGCTGCGCTGC  
80 90 100 110 120 130 140

150 160 170 180 190 200 X  
ACGAGGCTACACCCACGATGACCGTCCCCCGGCGATCGAGTCCCGGTGAGAGCGACGACAT  
|||||  
ACGAGGCTACACCCACGATGATGTTCCGGCGGCGATCGAGTCCCGGTGAGAGCGACGACAT  
150 160 170 180 190 200 X

19. US-09-697-123B-15 (1-208)  
US-09-697-123B-9 Sequence 9, Application US/09697123B

Initial Score = 148 Optimized Score = 175 Significance = -0.20  
Residue Identity = 84% Matches = 189 Mismatches = 19  
Gaps = 15 Conservative Substitutions = 0

X 10 20 30 40 50 60 70  
TCAAGAGAAAGCGCTACGACCTGGCCGGGTGGCGGCTCAAGGTCAACAGAAAGCTGCGCTGCACGCG  
|||||



```
|||||
TCAAGAGAGGCTACGACCTGGCCCGGCTGACAGTCAAGTCAAGAGCTGGTCTGCAAGCCG
X 10 20 30 40 50 60 70
80 90 100 110 120 130 140
GTGAGCCGATACAGCTGACGCTGACCGAGAGAGAGTGTGCGCCACATCGATGCTGGTGGAGCCG
|||||
CGGAGCCGATACAGCTGACGCTGACCGAGAGAGAGTGTGCGCCACATCGATGCTGGTGGAGCCG
|||||
CGGAGCCGATACAGCTGACGCTGACCGAGAGAGAGTGTGCGCCACATCGATGCTGGTGGAGCCG
80 90 100 110 120 130 140
AC-----GAGGCTAGCCAC---GATGACCGTCCCGGCGCATCGAGTGGTGGAGACCG
|||||
ACCGCCCGCTAGGAGTGGCCGCGCGCTCATGACTGTCCCGCGGATCGAGGTGCGGCTGGAGACCG
150 160 170 180 190 200 210
X
ACGACAT
|||||
ACGACAT
220 X
```

20. US-09-697-123b-15 (1-208)  
US-09-697-123b-1 Sequence 13, Application US/09697123b

Initial Score = 143 Optimized Score = 172 Significance = -0.29  
Residue Identity = 81% Matches = 182 Mismatches = 26  
Gaps = 15 Conservative Substitutions = 0

```
X 10 20 30 40 50 60 70
TCAAGAGAGGCTACGACCTGGCCCGGCTGACAGTCAAGTCAAGAGAGTGTGCGCCACATCG
|||||
TCAAGAGAGGCTACGACCTGGCCCGGCTGACAGTCAAGTCAAGAGAGTGTGCGCCACATCG
10 20 30 40 50 60 70
GTCAGCCGATACAGCTGACGCTGACCGAGAGAGTGTGCGCCACATCGATGCTGGTGGAGCCG
|||||
ATCATCCGATACAGCTGACGCTGACCGAGAGAGTGTGCGCCACATCGATGCTGGTGGAGCCG
80 90 100 110 120 130 140
ACGAGG-----GTCAGCCGATGACGCTGACCGAGAGAGTGTGCGCCACATCGAGTGGAGAACCG
150 160 170 180 190 200 210
|||||
ACGACAT
|||||
ACGACAT
220 X
```

21. US-09-697-123b-15 (1-208)  
US-09-697-123b-7 Sequence 7, Application US/09697123b

Initial Score = 131 Optimized Score = 170 Significance = -0.53  
Residue Identity = 82% Matches = 177 Mismatches = 31  
Gaps = 6 Conservative Substitutions = 0

```
X 10 20 30 40 50 60 70
TCAAGAGAGGCTACGACCTGGCCCGGCTGACAGTCAAGAGAGAGTGTGCGCCACATCGAGTGG
|||||
TCAAGAGAGGCTACGACCTGGCCCGGCTGACAGTCAAGAGAGAGTGTGCGCCACATCGAGTGG
10 20 30 40 50 60 70
GTGAGC-----GATCAGCCGATGACGCTGACCGAGAGAGAGTGTGCGCCACATCGAGTGG
|||||
CCAAGCCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
80 90 100 110 120 130 140
```

```
140 150 160 170 180 190 200
GCCTGCAGAGGCTACGACCTGGCCCGGCTGACAGTCAAGTCAAGAGAGTGTGCGCCACATCG
|||||
GCCTGCAGAGGCTACGACCTGGCCCGGCTGACAGTCAAGTCAAGAGAGTGTGCGCCACATCG
150 160 170 180 190 200 210
22. US-09-697-123b-15 (1-208)  
US-09-697-123b-1 Sequence 18, Application US/09697123b
```

Initial Score = 117 Optimized Score = 162 Significance = -0.80  
Residue Identity = 79% Matches = 168 Mismatches = 40  
Gaps = 3 Conservative Substitutions = 0

```
X 10 20 30 40 50 60 70
TCAAGAGAGGCTACGACCTGGCCCGGCTGACAGTCAAGTCAAGAGAGTGTGCGCCACATCG
|||||
TCAAGAGAGGCTACGACCTGGCCCGGCTGACAGTCAAGTCAAGAGAGTGTGCGCCACATCG
10 20 30 40 50 60 70
GTGAGCC-----GATCAGCCGATGACGCTGACCGAGAGAGAGTGTGCGCCACATCGAGTGG
|||||
ACCGCCCGCTAGGAGTGGCCGCGCGCTCATGACTGTCCCGCGGATCGAGGTGCGGCTGGAGACCG
80 90 100 110 120 130 140
TCAAGAGAGGCTACGACCTGGCCCGGCTGACAGTCAAGTCAAGAGAGTGTGCGCCACATCG
|||||
TCAAGAGAGGCTACGACCTGGCCCGGCTGACAGTCAAGTCAAGAGAGTGTGCGCCACATCG
150 160 170 180 190 200 210
```

23. US-09-697-123b-15 (1-208)  
US-09-697-123b-2 Sequence 21, Application US/09697123b

Initial Score = 112 Optimized Score = 172 Significance = -0.90  
Residue Identity = 83% Matches = 179 Mismatches = 29  
Gaps = 6 Conservative Substitutions = 0

```
X 10 20 30 40 50 60 70
TCAAGAGAGGCTACGACCTGGCCCGGCTGACAGTCAAGTCAAGAGAGTGTGCGCCACATCG
|||||
TCAAGAGAGGCTACGACCTGGCCCGGCTGACAGTCAAGTCAAGAGAGTGTGCGCCACATCG
10 20 30 40 50 60 70
GTGAGCC-----GATCAGCCGATGACGCTGACCGAGAGAGAGTGTGCGCCACATCGAGTGG
|||||
CCAATCCGCTGACAGTGGACACCAACCTTCAAGAGAGAGTGTGCGCCACATCGAGTGGTGC
80 90 100 110 120 130 140
```

24. US-09-697-123b-15 (1-208)  
US-09-697-123b-1 Sequence 14, Application US/09697123b

Initial Score = 112 Optimized Score = 173 Significance = -0.90  
Residue Identity = 84% Matches = 180 Mismatches = 28  
Gaps = 6 Conservative Substitutions = 0

```
X 10 20 30 40 50 60 70
TCAAGAGAGGCTACGACCTGGCCCGGCTGACAGTCAAGTCAAGAGAGTGTGCGCCACATCG
|||||
TCAAGAGAGGCTACGACCTGGCCCGGCTGACAGTCAAGTCAAGAGAGTGTGCGCCACATCG
10 20 30 40 50 60 70
```

25. US-09-697-123B-15 (1-208)  
US-09-697-123B-2 Sequence 25, Application US/09697123B

Initial Score	=	19	Optimized Score	=	19	Significance	=	-2.73
Residue Identity	=	100%	Matches	=	19	Mismatches	=	0
Gaps	=	0	Conservative Substitutions	=	0			0

> O <  
O/I/O Intelligenetics  
> O <

FastDB - Fast Pairwise Comparison of Sequences  
Release 5.4

Results file us-09-697-123b-14.res made by shanley on Wed 13 Nov 102 14:07:43-PST.

Query sequence being compared: US-09-697-123B-14 (1-214)  
Number of sequences searched: 26  
Number of scores above cutoff: 25

Results of the initial comparison of US-09-697-123B-14 (1-214) with:  
File : US09697123B.seq

```

100-
N -
U -
M -
B -
E -
R -
O -
F 10-
S -
S -
E 5-
O -
U -
E -
N -
C -
S 0-
SCORE 0 | 24 | 48 | 71 | 95 | 119 | 143 | 166 | 190 | 214 |
STDEV -2 | -1 | -1 | -1 | -1 | -1 | -1 | -1 | -1 | -1 |
PARAMETERS
Similarity matrix Unlary K-tuple
Mismatch penalty 1 Joining penalty 4
Gap penalty 5.00 Window size 30
Gap size penalty 0.33 207
Cutoff score 12
Randomization group 0
SEARCH STATISTICS
Scores: Mean 113 Median 113 Standard Deviation 40.53
Times: CPU 00:00:00.00 Total Elapsed 00:00:00.00
Number of residues: 5077
Number of sequences searched: 26
Number of scores above cutoff: 25

```

The scores below are sorted by initial score.  
Significance is calculated based on initial score.

A 100% identical sequence to the query sequence was found:

Sequence Name	Description	Length	Score	Int. Opt.	Sig.	Frame
US-09-697-123B-14	Sequence 14, Application	214	214	214	2.49	0

1. US-09-697-123B-1 Sequence 14, Application 214 214 214 2.49 0  
The list of other best scores is:

Sequence Name	Description	Length	Score	Int. Opt.	Sig.	Frame
2. US-09-697-123B-2 Sequence 7, Application	214	172	172	1.46	0	
3. US-09-697-123B-2 Sequence 21, Application	214	171	171	1.43	0	
4. US-09-697-123B-1 Sequence 1, Application	208	136	169	0.57	0	
5. US-09-697-123B-2 Sequence 23, Application	208	132	166	0.47	0	
6. US-09-697-123B-1 Sequence 11, Application	208	129	166	0.39	0	
7. US-09-697-123B-1 Sequence 10, Application	208	126	163	0.32	0	
8. US-09-697-123B-2 Sequence 20, Application	205	122	163	0.22	0	
9. US-09-697-123B-1 Sequence 12, Application	207	120	167	0.17	0	
10. US-09-697-123B-8 Sequence 8, Application	208	118	168	0.12	0	
11. US-09-697-123B-1 Sequence 19, Application	208	116	168	0.07	0	
12. US-09-697-123B-1 Sequence 18, Application	211	115	166	0.05	0	
13. US-09-697-123B-1 Sequence 15, Application	208	112	173	-0.02	0	
14. US-09-697-123B-3 Sequence 3, Application	208	112	168	-0.02	0	
15. US-09-697-123B-2 Sequence 2, Application	208	112	170	-0.02	0	
16. US-09-697-123B-2 Sequence 16, Application	208	110	168	-0.07	0	
17. US-09-697-123B-6 Sequence 6, Application	208	110	168	-0.07	0	
18. US-09-697-123B-5 Sequence 5, Application	208	110	168	-0.07	0	
19. US-09-697-123B-2 Sequence 24, Application	208	109	171	-0.10	0	
20. US-09-697-123B-2 Sequence 22, Application	208	109	167	-0.10	0	
21. US-09-697-123B-1 Sequence 17, Application	208	109	169	-0.10	0	
22. US-09-697-123B-4 Sequence 4, Application	207	97	161	-0.39	0	
23. US-09-697-123B-1 Sequence 13, Application	223	83	152	-0.74	0	
24. US-09-697-123B-9 Sequence 9, Application	223	81	156	-0.79	0	
25. US-09-697-123B-2 Sequence 25, Application	19	19	19	-2.32	0	

1. US-09-697-123B-14 (1-214)  
US-09-697-123B-1 Sequence 14, Application US/09697123B  
Initial Score = 214 Optimized Score = 214 Significance = 2.49  
Residue Identity = 100% Matches = 214 Mismatches = 0  
Gaps = 0 Conservative Substitutions = 0

```

X 10 20 30 40 50 60 70
TCAAGGAGAGCGCTACGACCTGCGCCAGCGGTTGCCGCTTACAGTCAAGTCAAGACCTGCGCGCGG
|||||
TCMAAGAGAGAGCGCTACGACCTGCGCCAGCGGTTGCCGCTTACAGTCAAGTCAAGACCTGCGCGCGG
X 10 20 30 40 50 60 70

```

```

80 90 100 110 120 130 140
CCGAGTCGCGCGCTACGACCTGCGCCAGCGGTTGCCGCTTACAGTCAAGTCAAGACCTGCGCGCGG
|||||
CCGAGTCGCGCGCTACGACCTGCGCCAGCGGTTGCCGCTTACAGTCAAGTCAAGACCTGCGCGCGG
80 90 100 110 120 130 140

```

```

150 160 170 180 190 200 210
GCCGTCAGAGAGCGCTACGACCTGCGCCAGCGGTTGCCGCTTACAGTCAAGTCAAGACCTGCGCGG
|||||
GCCGTCAGAGAGCGCTACGACCTGCGCCAGCGGTTGCCGCTTACAGTCAAGTCAAGACCTGCGCGG
150 160 170 180 190 200 210

```

2. US-09-697-123B-14 (1-214)  
US-09-697-123B-7 Sequence 7, Application US/09697123B

Initial Score = 172 Optimized Score = 172 Significance = 1.46  
Residue Identity = 80% Matches = 172 Mismatches = 42  
Gaps = 0 Conservative Substitutions = 0

```

X 10 20 30 40 50 60 70
TCAAGGAGAGCGCTACGACCTGCGCCAGCGGTTGCCGCTTACAGTCAAGTCAAGACCTGCGCGG
|||||
TCAAGGAGAGCGCTACGACCTGCGCCAGCGGTTGCCGCTTACAGTCAAGTCAAGACCTGCGCGG
63 64 65 66 67 68 69

```

Initial Score	=	132	Optimized Score	=	169	Significance	=	0.47
Residue Identity	=	82%	Matches	=	176	Mismatches	=	32
Gaps	=	6	Conservative Substitutions	=			=	0

150 160 170 180 190 200 210 X  
CCCTGCACGAGGCGCGAGCAAGATGACGTTCCCGCGCGCTCAGAGTGCCGGTGGAGACCGACGACAT  
CTTGGACGAGGGGCGAGACCGGATACCGTCCGGGCGGCTGCGAGTGTCGCCGCTCGAGACCGACGACAT  
150 160 170 180 190 200 X

8. US-09-697-123B-14 (1-214)

CCGAGTGTGGCCCTACCCGCTTCGACCAACGCTGACCCGAAGCGGATGTGTCGTCGCCACATCGAGTACCGGTG  
 80 90 100 110 120 130 140  
 111 111 111 111 111 111 111  
 GTGAGCGG-----ATCACACAGCTGCACCGCTGACCCGAGGAAGACGTCGTCGCCACATCGAGTACCGGTG

14. US-09-697-123B-14 (1-214)  
US-09-697-123B-3 Sequence 3, Application US/09697123B

Initial Score = 112 Optimized Score = 168 Significance = -0.02  
Residue Identity = 81% Matches = 175 Mismatches = 33  
Gaps = 6 Conservative Substitutions = 0

150 160 170 180 190 200 210 X  
GCCGTGCAGAGGCGCCAGCAACGATATGACGTTCCCGCGCGCTGAGGTGCCCGGTGAGACCGACGACAT  
|||||  
GCCGTGCAGAGGCGCCAGCAACGATATGACGTTCCCGCGCGCTGAGGTGCCCGGTGAGACCGACGACAT  
140 150 160 170 180 190 200 X

150 160 170 180 190 200 210 X  
CCGAGTGCAGGCGCCAGCAACGATATGACGTTCCCGCGCGCTGAGGTGCCCGGTGAGACCGACGACAT  
|||||  
CCGAGTGCAGGCGCCAGCAACGATATGACGTTCCCGCGCGCTGAGGTGCCCGGTGAGACCGACGACAT  
140 150 160 170 180 190 200 X

150 160 170 180 190 200 210 X  
GCTTGCACGAGGCGCCAGCAACGATATGACGTTCCCGCGCGCTGAGGTGCCCGGTGAGACCGACGACAT  
|||||  
GCTTGCACGAGGCGCCAGCAACGATATGACGTTCCCGCGCGCTGAGGTGCCCGGTGAGACCGACGACAT  
140 150 160 170 180 190 200 X

15. US-09-697-123B-14 (1-214)  
US-09-697-123B-2 Sequence 2, Application US/09697123B

Initial Score = 112 Optimized Score = 170 Significance = -0.02  
Residue Identity = 82% Matches = 177 Mismatches = 31  
Gaps = 6 Conservative Substitutions = 0

150 160 170 180 190 200 210 X  
TCAAGGAGAAAGCGCTACGACCTGCGCCAGCAACGATATGACGTTCCCGCGCGCTGAGGTGCCCGGTGAGACCGACGACAT  
|||||  
TCAAGGAGAAAGCGCTACGACCTGCGCCAGCAACGATATGACGTTCCCGCGCGCTGAGGTGCCCGGTGAGACCGACGACAT  
140 150 160 170 180 190 200 X

150 160 170 180 190 200 210 X  
CCGAGTGCAGGCGCCAGCAACGATATGACGTTCCCGCGCGCTGAGGTGCCCGGTGAGACCGACGACAT  
|||||  
CCGAGTGCAGGCGCCAGCAACGATATGACGTTCCCGCGCGCTGAGGTGCCCGGTGAGACCGACGACAT  
140 150 160 170 180 190 200 X

150 160 170 180 190 200 210 X  
GCTTGCACGAGGCGCCAGCAACGATATGACGTTCCCGCGCGCTGAGGTGCCCGGTGAGACCGACGACAT  
|||||  
GCTTGCACGAGGCGCCAGCAACGATATGACGTTCCCGCGCGCTGAGGTGCCCGGTGAGACCGACGACAT  
140 150 160 170 180 190 200 X

16. US-09-697-123B-14 (1-214)  
US-09-697-123B-1 Sequence 16, Application US/09697123B

Initial Score = 110 Optimized Score = 168 Significance = -0.07  
Residue Identity = 81% Matches = 175 Mismatches = 33  
Gaps = 6 Conservative Substitutions = 0

150 160 170 180 190 200 210 X  
TCAAGGAGAAAGCGCTACGACCTGCGCCAGCAACGATATGACGTTCCCGCGCGCTGAGGTGCCCGGTGAGACCGACGACAT  
|||||  
TCAAGGAGAAAGCGCTACGACCTGCGCCAGCAACGATATGACGTTCCCGCGCGCTGAGGTGCCCGGTGAGACCGACGACAT  
140 150 160 170 180 190 200 X

150 160 170 180 190 200 210 X  
GCCGTGCAGAGGCGCCAGCAACGATATGACGTTCCCGCGCGCTGAGGTGCCCGGTGAGACCGACGACAT  
|||||  
GCCGTGCAGAGGCGCCAGCAACGATATGACGTTCCCGCGCGCTGAGGTGCCCGGTGAGACCGACGACAT  
140 150 160 170 180 190 200 X

150 160 170 180 190 200 210 X  
GCCGTGCAGAGGCGCCAGCAACGATATGACGTTCCCGCGCGCTGAGGTGCCCGGTGAGACCGACGACAT  
|||||  
GCCGTGCAGAGGCGCCAGCAACGATATGACGTTCCCGCGCGCTGAGGTGCCCGGTGAGACCGACGACAT  
140 150 160 170 180 190 200 X

17. US-09-697-123B-14 (1-214)  
US-09-697-123B-6 Sequence 6, Application US/09697123B

Initial Score = 110 Optimized Score = 168 Significance = -0.07  
Residue Identity = 81% Matches = 175 Mismatches = 33  
Gaps = 6 Conservative Substitutions = 0

150 160 170 180 190 200 210 X  
TCAAGGAGAAAGCGCTACGACCTGCGCCAGCAACGATATGACGTTCCCGCGCGCTGAGGTGCCCGGTGAGACCGACGACAT  
|||||  
TCAAGGAGAAAGCGCTACGACCTGCGCCAGCAACGATATGACGTTCCCGCGCGCTGAGGTGCCCGGTGAGACCGACGACAT  
140 150 160 170 180 190 200 X

150 160 170 180 190 200 210 X  
CCGAGTGCAGGCGCCAGCAACGATATGACGTTCCCGCGCGCTGAGGTGCCCGGTGAGACCGACGACAT  
|||||  
CCGAGTGCAGGCGCCAGCAACGATATGACGTTCCCGCGCGCTGAGGTGCCCGGTGAGACCGACGACAT  
140 150 160 170 180 190 200 X

150 160 170 180 190 200 210 X  
GCTTGCACGAGGCGCCAGCAACGATATGACGTTCCCGCGCGCTGAGGTGCCCGGTGAGACCGACGACAT  
|||||  
GCTTGCACGAGGCGCCAGCAACGATATGACGTTCCCGCGCGCTGAGGTGCCCGGTGAGACCGACGACAT  
140 150 160 170 180 190 200 X

18. US-09-697-123B-14 (1-214)  
US-09-697-123B-5 Sequence 5, Application US/09697123B

Initial Score = 110 Optimized Score = 168 Significance = -0.07  
Residue Identity = 81% Matches = 175 Mismatches = 33  
Gaps = 6 Conservative Substitutions = 0

150 160 170 180 190 200 210 X  
TCAAGGAGAAAGCGCTACGACCTGCGCCAGCAACGATATGACGTTCCCGCGCGCTGAGGTGCCCGGTGAGACCGACGACAT  
|||||  
TCAAGGAGAAAGCGCTACGACCTGCGCCAGCAACGATATGACGTTCCCGCGCGCTGAGGTGCCCGGTGAGACCGACGACAT  
140 150 160 170 180 190 200 X

150 160 170 180 190 200 210 X  
CCGAGTGCAGGCGCCAGCAACGATATGACGTTCCCGCGCGCTGAGGTGCCCGGTGAGACCGACGACAT  
|||||  
CCGAGTGCAGGCGCCAGCAACGATATGACGTTCCCGCGCGCTGAGGTGCCCGGTGAGACCGACGACAT  
140 150 160 170 180 190 200 X

150 160 170 180 190 200 210 X  
GCTTGCACGAGGCGCCAGCAACGATATGACGTTCCCGCGCGCTGAGGTGCCCGGTGAGACCGACGACAT  
|||||  
GCTTGCACGAGGCGCCAGCAACGATATGACGTTCCCGCGCGCTGAGGTGCCCGGTGAGACCGACGACAT  
140 150 160 170 180 190 200 X

19. US-09-697-123B-14 (1-214)  
US-09-697-123B-2 Sequence 24, Application US/09697123B

Initial Score = 109 Optimized Score = 171 Significance = -0.10  
Residue Identity = 83% Matches = 178 Mismatches = 30  
Gaps = 6 Conservative Substitutions = 0

10 20 30 40 50 60 70  
X  
TCAAGAGAAAGGCGTACGACCTGGCCAGAGGTTGGCGGTTCACAGTCAACAAAGTGGGCTGCCGCGG  
|||||  
TCAAGAGAAAGGCGTACGACCTGGCCAGAGGTTGGCGGTTCACAGTCAACAAAGTGGGCTGCCGCGG  
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140  
CCGAGTGGCCGCTACCCGCTTCGACACCGCTGACCGAAGCGGATGTGGCGCCACCATCGAGTACCTGGTGC  
|||  
AGATGGC-----CCAAACACACACACCTCGACCGAAGAGAGCTGTCTGCCCATCTGAATACCTGGTGC  
80 90 100 110 120 130

150 160 170 180 190 200 210  
X  
GCCTGACAGAGGCGCCAGGCAACGATGACGGTCCCGGCGCGCTGCGAGGTCGGTGGAGACCGAGACAT  
|||||  
GCTTGCACAGAGGCGCCAGGCAACGATGACGGTCCCGGCGCGCTGCGAGGTCGGTGGAGACCGAGACAT  
140 150 160 170 180 190 200 X

## 20. US-09-697-123B-14 (1-214)

US-09-697-123B-2 sequence 22, Application US/09697123B

Initial Score = 109 Optimized Score = 167 Significance = -0.10  
Residue Identity = 81% Matches = 174 Mismatches = 34  
Gaps = 6 Conservative Substitutions = 0

10 20 30 40 50 60 70  
X  
TCAAGAGAAAGGCGTACGACCTGGCCAGAGGTTGGCGGTTCACAGTCAACAAAGTGGGCTGCCGCGG  
|||||  
TCAAGAGAAAGGCGTACGACCTGGCCAGAGGTTGGCGGTTCACAGTCAACAAAGTGGGCTGCCGCGG  
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140  
CCGAGTGGCCGCTACCCGCTTCGACACCGCTGACCGAAGCGGATGTGGCGCCACCATCGAGTACCTGGTGC  
|||||  
GCGAGC-----CATCAGCTGCTGCGACGCTGACCAAGAACAGCTGCGCACCATCTGAATATCTGGTGC  
80 90 100 110 120 130

150 160 170 180 190 200 210  
X  
GCCTGACAGAGGCGCCAGGCAACGATGACGGTCCCGGCGCGCTGCGAGGTCGGTGGAGACCGAGACAT  
|||||  
GCTTGCACAGAGGCGCCAGGCAACGATGACGGTCCCGGCGCGCTGCGAGGTCGGTGGAGACCGAGACAT  
140 150 160 170 180 190 200 X

## 21. US-09-697-123B-14 (1-214)

US-09-697-123B-1 sequence 17, Application US/09697123B

Initial Score = 109 Optimized Score = 169 Significance = -0.10  
Residue Identity = 82% Matches = 176 Mismatches = 32  
Gaps = 6 Conservative Substitutions = 0

10 20 30 40 50 60 70  
X  
TCAAGAGAAAGGCGTACGACCTGGCCAGAGGTTGGCGGTTCACAGTCAACAAAGTGGGCTGCCGCGG  
|||||  
TCAAGAGAAAGGCGTACGACCTGGCCAGAGGTTGGCGGTTCACAGTCAACAAAGTGGGCTGCCGCGG  
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140  
CCGAGTGGCCGCTACCCGCTTCGACACCGCTGACCGAAGCGGATGTGGCGCCACCATCGAGTACCTGGTGC  
|||||  
-----GCTGCCGCTACGACCTGCTGACCGAAGAGAGCTGCTGCCCAACCATCGAGTACCTGGTGC  
80 90 100 110 120 130

150 160 170 180 190 200 210  
X  
GCCTGACAGAGGCGCCAGGCAACGATGACGGTCCCGGCGCGCTGCGAGGTCGGTGGAGACCGAGACAT  
|||||  
GCTTGCACAGAGGCGCCAGGCAACGATGACGGTCCCGGCGCGCTGCGAGGTCGGTGGAGACCGAGACAT  
140 150 160 170 180 190 200 X

Initial Score = 97 Optimized Score = 161 Significance = -0.39  
Residue Identity = 80% Matches = 173 Mismatches = 34  
Gaps = 7 Conservative Substitutions = 0

10 20 30 40 50 60 70  
X  
TCAAGAGAAAGGCGTACGACCTGGCCAGAGGTTGGCGGTTCACAGTCAACAAAGTGGGCTGCCGCGG  
|||||  
TCAAGAGAAAGGCGTACGACCTGGCCAGAGGTTGGCGGTTCACAGTCAACAAAGTGGGCTGCCGCGG  
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140  
CCGAGTGGCCGCTACCCGCTTCGACACCGCTGACCGAAGCGGATGTGGCGCCACCATCGAGTACCTGGTGC  
|||||  
GCGATCG-----ATCAGCAGCTGCGACGCTGACCGAAGAGAGCTGTCTGCCCATCTGAATACCTGGTGC  
80 90 100 110 120 130

150 160 170 180 190 200 210  
X  
GCCTGACAGAGGCGCCAGGCAACGATGACGGTCCCGGCGCGCTGCGAGGTCGGTGGAGACCGAGACAT  
|||||  
GCTTGCACAGAGGCGCCAGGCAACGATGACGGTCCCGGCGCGCTGCGAGGTCGGTGGAGACCGAGACAT  
140 150 160 170 180 190 200 X

## 23. US-09-697-123B-14 (1-214)

US-09-697-123B-1 sequence 13, Application US/09697123B

Initial Score = 83 Optimized Score = 152 Significance = -0.74  
Residue Identity = 75% Matches = 173 Mismatches = 35  
Gaps = 21 Conservative Substitutions = 0

10 20 30 40 50 60 70  
X  
TCAAGAGAAAGGCGTACGACCTGGCCAGAGGTTGGCGGTTCACAGTCAACAAAGTGGGCTGCCGCGG  
|||||  
TCAAGAGAAAGGCGTACGACCTGGCCAGAGGTTGGCGGTTCACAGTCAACAAAGTGGGCTGCCGCGG  
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140  
CCGAGTGGCCGCTACCCGCTTCGACACCGCTGACCGAAGCGGATGTGGCGCCACCATCGAGTACCTGGTGC  
|||||  
ATCATCG-----ATCAGCAGCTGCGACGCTGACCGAAGAGAGCTGTCTGCCCATCTGAATACCTGGTGC  
80 90 100 110 120 130

150 160 170 180 190 200 210  
X  
GCCTGACAGAGGCGCCAGGCAACGATGACGGTCCCGGCGCGCTGCGAGGTCGGTGGAGACCGAGACAT  
|||||  
GCTTGCACAGAGGCGCCAGGCAACGATGACGGTCCCGGCGCGCTGCGAGGTCGGTGGAGACCGAGACAT  
140 150 160 170 180 190 200 X

210 X  
AGACGAGAGACAT  
AAACCCAGACAT  
220 X

## 24. US-09-697-123B-14 (1-214)

US-09-697-123B-9 sequence 9, Application US/09697123B

Initial Score = 81 Optimized Score = 156 Significance = -0.79  
Residue Identity = 75% Matches = 172 Mismatches = 36  
Gaps = 21 Conservative Substitutions = 0

10 20 30 40 50 60 70  
X  
TCAAGAGAAAGGCGTACGACCTGGCCAGAGGTTGGCGGTTCACAGTCAACAAAGTGGGCTGCCGCGG  
|||||  
TCAAGAGAAAGGCGTACGACCTGGCCAGAGGTTGGCGGTTCACAGTCAACAAAGTGGGCTGCCGCGG  
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140  
CCGAGTGGCCGCTACCCGCTTCGACACCGCTGACCGAAGCGGATGTGGCGCCACCATCGAGTACCTGGTGC  
|||||  
GCGAGCG-----ATCAGCAGCTGCGACGCTGACCGAAGAGAGCTGTCTGCCCATCTGAATACCTGGTGC  
80 90 100 110 120 130

25. US-09-697-123B-14 (1-214)  
US-09-697-123B-2 Sequence 25, Application US/09697123B

Initial Score	=	19	Optimized Score	=	19	Significance	=	-2.32
Residue Identity	=	100%	Matches	=	19	Mismatches	=	0
Gaps	=	0	Conservative Substitutions	=	0			

X 10 X  
TCAAGGAGAAAGCGCTACGA  
|||||  
TCAAGGAGAAAGCGCTACGA  
X 10 X



2. US-09-697-123B-13 (1-223)  
US-09-697-123B-9 Sequence 9, Application US/09697123B

Initial Score	=	195	Optimized Score	=	195	Significance	=	1.54
---------------	---	-----	-----------------	---	-----	--------------	---	------

150  
160  
170  
180  
190  
200  
210

ACGAGC-----GTACCCACGATGACCGTCCCGGCGCATCGAGTGGCGGTGAGACCG  
150 160 170 180 190 200  
220 X  
ACGACAT  
|||||  
ACGACAT  
X

7. US-09-697-123B-13 (1-223)  
US-09-697-123B-3 Sequence 3, Application US/09697123B

Initial Score = 142 Optimized Score = 167 Significance = 0.34  
Residue Identity = 79% Matches = 177 Mismatches = 31  
Gaps = 15 Conservative Substitutions = 0

X 10 20 30 40 50 60 70  
TCAAGAGAGAGCGCTACGACGTCGGCCCGCGGTACAAAGGTCAAGAAGAGTGGGCTGAAACCG  
|||||  
TCAAGAGAGAGCGCTACGACGTCGGCCCGCGGTACAAAGGTCAAGAAGAGTGGGCTGAAACCG  
X 10 20 30 40 50 60 70  
ATCATCCGATCACCAACGACGCTGACCGAAGAGACGTCTGCGCCACATCGAGTACCTGTTCCGCTGC  
|||||  
CGCATCCGATCACCAACGACGCTGACCGAAGAGACGTCTGCGCCACATCGAGTACCTGTTCCGCTGC  
80 90 100 110 120 130 140  
ACCAGCGCTCTCAGGGTGGCCAGGCCCGCTTATGACGTGTCGCCGCGGGTGCAGGTGCGGTGAACCG  
|||||  
ACGAGGCTCAGCA-----CAGCATGACCGTTCCGGGCGGCACCGAGGTTCGGGTGGAACCG  
150 160 170 180 190 200

220 X  
ACGACAT  
|||||  
ACGACAT  
X

8. US-09-697-123B-13 (1-223)  
US-09-697-123B-1 Sequence 1, Application US/09697123B

Initial Score = 142 Optimized Score = 168 Significance = 0.34  
Residue Identity = 79% Matches = 178 Mismatches = 30  
Gaps = 15 Conservative Substitutions = 0

X 10 20 30 40 50 60 70  
TCAAGAGAGAGCGCTACGACGTCGGCCCGCGGTACAAAGGTCAAGAAGAGTGGGCTGAAACCG  
|||||  
TCAAGAGAGAGCGCTACGACGTCGGCCCGCGGTACAAAGGTCAAGAAGAGTGGGCTGAAACCG  
X 10 20 30 40 50 60 70

ATCATCCGATCACCAACGACGCTGACCGAAGAGACGTCTGCGCCACATCGAGTACCTGTTCCGCTGC  
|||||  
GCGATCCGATCACCAACGACGCTGACCGAAGAGACGTCTGCGCCACATCGAGTACCTGTTCCGCTGC  
80 90 100 110 120 130 140

150 160 170 180 190 200 210  
ACCAAGCGCTCTCAGGGTGGCCAGGCCCGCTTATGACTGTCCCGCGGGTGCAGGTGCGGTGAACCG  
|||||  
ACGAGGCGCAACA-----CAGCATGACCGTTCCGGGCGGCACCGAGGTGCGGTGGAACCG  
150 160 170 180 190 200

220 X  
ACGACAT  
|||||  
ACGACAT  
X

9. US-09-697-123B-13 (1-223)  
US-09-697-123B-1 Sequence 11, Application US/09697123B

Initial Score = 141 Optimized Score = 167 Significance = 0.32  
Residue Identity = 79% Matches = 177 Mismatches = 31  
Gaps = 15 Conservative Substitutions = 0

X 10 20 30 40 50 60 70  
TCAAGAGAGAGCGCTACGACGTCGGCCCGCGGTACAAAGGTCAAGAAGAGTGGGCTGAAACCG  
|||||  
TCAAGAGAGAGCGCTACGACGTCGGCCCGCGGTACAAAGGTCAAGAAGAGTGGGCTGAAACCG  
X 10 20 30 40 50 60 70

ATCATCCGATCACCAACGACGCTGACCGAAGAGACGTCTGCGCCACATCGAGTACCTGTTCCGCTGC  
|||||  
GCGATCCGATCACCAACGACGCTGACCGAAGAGACGTCTGCGCCACATCGAGTACCTGTTCCGCTGC  
80 90 100 110 120 130 140

150 160 170 180 190 200 210  
ACCAGCGCTCTCAGGGTGGCCAGGCCCGCTTATGACTGTCCCGCGGGTGCAGGTGCGGTGAACCG  
|||||  
ACGAGGCTCAGCA-----CAGCATGACCGTTCCGGGCGGCACCGAGGTTCGGGTGGAACCG  
150 160 170 180 190 200

220 X  
ACGACAT  
|||||  
ACGACAT  
X

10. US-09-697-123B-13 (1-223)  
US-09-697-123B-2 Sequence 24, Application US/09697123B

Initial Score = 140 Optimized Score = 168 Significance = 0.29  
Residue Identity = 79% Matches = 178 Mismatches = 30  
Gaps = 15 Conservative Substitutions = 0

X 10 20 30 40 50 60 70  
TCAAGAGAGAGCGCTACGACGTCGGCCCGCGGTACAAAGGTCAAGAAGAGTGGGCTGAAACCG  
|||||  
TCAAGAGAGAGCGCTACGACGTCGGCCCGCGGTACAAAGGTCAAGAAGAGTGGGCTGAAACCG  
X 10 20 30 40 50 60 70

ATCATCCGATCACCAACGACGCTGACCGAAGAGACGTCTGCGCCACATCGAGTACCTGTTCCGCTGC  
|||||  
AGATGCGCAACCAACGACGCTGACCGAAGAGACGTCTGCGCCACATCGAGTACCTGTTCCGCTGC  
80 90 100 110 120 130 140

150 160 170 180 190 200 210  
ACCAAGCGCTCTCAGGGTGGCCAGGCCCGCTTATGACTGTCCCGCGGGTGCAGGTGCGGTGAACCG  
|||||  
ACGAGG-----GCGATGACCGTGAAGTCCCGGCTGCGAGTCCCGGTGGAACCG  
150 160 170 180 190 200

220 X  
ACGACAT  
|||||  
ACGACAT  
X

11. US-09-697-123B-13 (1-223)  
US-09-697-123B-1 Sequence 10, Application US/09697123B

Initial Score = 140 Optimized Score = 166 Significance = 0.29  
Residue Identity = 78% Matches = 176 Mismatches = 32  
Gaps = 15 Conservative Substitutions = 0

X 10 20 30 40 50 60 70

[illegible]

150 160 170 180 190 200 210  
 ACCACCCCTCTCAGGGTGGCCAGGCCCGCTTATGACGTCTCCCGCGGGGTGAGGTGGCGGTGAACCG  
 AC-----GAGGTCACACAC--GATGACGGTTCGGGGGGGCGTGCAGTCCGCTGGAACCG  
 150 160 170 180 190 200  
 220 X

✕

US-09-697-123B-5 Sequence

5

X



X

US-09-697-123B-2 Sequence

3

x



✕

US-09-697-123B-2 Sequence

## Gaps

**X**

Q

X

US-09-697-123B-4 Sequence

6

x

aa

X

05-09-697-123B-1 Sequence

6

Y

140 150 160 170 180 190 200 210

GCCTGACACACGCTCTCAGGGTGGCCAGGCCCCCGTTATGACTGTCCCGCCGGTCGAGTGCCTGGTGG

220 X  
AAACCGACGACAT  
| |||||  
ATGTGACGACAT  
210 X

25. US-09-697-123B-13 (1-223)  
US-09-697-123B-2 Sequence 25, Application US/09697123B  
Initial Score = 19 Optimized Score = 19 Significance = -2.44  
Residue Identity = 100% Matches = 19 Mismatches = 0  
Gaps = 0 Conservative Substitutions = 0  
X 10 X  
TCACGAGAGAGCGCTACGA  
|||||  
TCACGAGAGAGCGCTACGA  
X 10 X

---





x	10	20	30	40	50	60	70
---	----	----	----	----	----	----	----

8. US-09-697-123B-12 (1-207)  
US-09-697-123B-1 Sequence 16, Application US/09697123B



10 20 30 40 50 60 70  
 CCAGGAGAGCGCTACGACCTGTC-GCGTCGGCCGTTACAAGTCAACAAAGAAGCTCGGCTCGA

```
|||||
TCAAGAGAACCGCTACGACCTGGCCGAGGTGGCCGTTCAACAGTCAACAAAGAACTGGGGCTGCCGCGG
X 10 20 30 40 50 60 70
GCGAGCCG-----ATCACCAGTTGACCGCTGACCGCAAGAGAGATGTCGCCACCATGATACCTGGTTC
|||||
CCGAGTGGCGCCCTGACCGCTCGACCGCTGACCGAAGGGGATGTCGTCGCCACCATGATACCTGGTTC
80 90 100 110 120 130 140
GCGAGCGC-----ATCACCAGTTGACCGCTGACCGCAAGAGAGATGTCGCCACCATGATACCTGGTTC
|||||
GCTGTCACAGAGCGCCACGATGATGCTCCCGCGCGCTGAGGTGCGCGTGGAGACCGACGACAT
150 160 170 180 190 200 210 X
```

20. US-09-697-123b-12 (1-207)  
US-09-697-123b-1 Sequence 13, Application US/09697123B

Initial Score = 112 Optimized Score = 160 Significance = -0.51  
Residue Identity = 80% Matches = 179 Mismatches = 28  
Gaps = 16 Conservative Substitutions = 0

```
X 10 20 30 40 50 60 70
TCAAGAGAGAGCGCTACGACCTGG-TGCGCTGCGCCGCTTACAAAGTCAACAAAGAAAGCTGCTGAACTG
|||||
TCAAGAGAGAGCGCTACGACCTGG-TGCGCTGCGCCGCTTACAAAGTCAACAAAGAAAGCTGCTGAACTG
X 10 20 30 40 50 60 70
GCGAGCGC-----ATCACCAGTTGACCGCTGACCGCAAGAGAGATGTCGCCACCATGATACCTGGTTC
|||||
ATCATCGATCCACACCGACGCTGACCGCAAGAGAGATGTCGCCACCATGATACCTGGTTC
80 90 100 110 120 130 140
AC-----GAGGGCCAGAACAC--GATGACCGCTTCCCGCGCACCGAGAGTGGAGACCG
|||||
ACACCGCTCTCTCAGGTGGCCAGAGCGCCGCTTATGACTGTCCCGCGCGGTGCGGTGCGAGAACCG
150 160 170 180 190 200 210
```

X  
ACGACAT  
|||||  
ACGACAT  
220 X

21. US-09-697-123b-12 (1-207)  
US-09-697-123b-9 Sequence 9, Application US/09697123B

Initial Score = 112 Optimized Score = 164 Significance = -0.51  
Residue Identity = 80% Matches = 179 Mismatches = 28  
Gaps = 16 Conservative Substitutions = 0

```
X 10 20 30 40 50 60 70
TCAAGAGAGAGCGCTACGACCTGG-TGCGCTGCGCCGCTTACAAAGTCAACAAAGCTGCTGAACTG
|||||
TCAAGAGAGAGCGCTACGACCTGG-TGCGCTGCGCCGCTTACAAAGTCAACAAAGCTGCTGAACTG
X 10 20 30 40 50 60 70
GCGAGCGC-----ATCACCAGTTGACCGCTGACCGCAAGAGAGATGTCGCCACCATGATACCTGGTTC
|||||
GCGAGCGC-----ATCACCAGTTGACCGCTGACCGCAAGAGAGATGTCGCCACCATGATACCTGGTTC
80 90 100 110 120 130 140
```

```
ACGAGGCGCAG-----CCACGATGACCGTTCCCGCGCGACCCAGAGTGGCGTGGAGACCG
|||||
ACGAGGCGCAG-----CCACGATGACCGTTCCCGCGCGACCCAGAGTGGCGTGGAGACCG
150 160 170 180 190 200 210
```

X  
ACGACAT  
|||||  
ACGACAT  
220 X

22. US-09-697-123b-12 (1-207)  
US-09-697-123b-2 Sequence 21, Application US/09697123B

Initial Score = 106 Optimized Score = 158 Significance = -0.64  
Residue Identity = 79% Matches = 170 Mismatches = 37  
Gaps = 7 Conservative Substitutions = 0

```
X 10 20 30 40 50 60 70
TCAAGAGAGAGCGCTACGACCTGG-TGCGCTGCGCCGCTTACAAAGTCAACAAAGAAAGCTGCTGAACTG
|||||
TCAAGAGAGAGCGCTACGACCTGG-TGCGCTGCGCCGCTTACAAAGTCAACAAAGAAAGCTGCTGAACTG
X 10 20 30 40 50 60 70
```

```
GCGAGCC-----GATCACCAGTTGACCGCTGACCGCAAGAGAGATGTCGCCACCATGATACCTGGTTC
|||||
CCATCGCGCTCAGGTGACCGACACCGCTTACAAAGTCAACAAAGAAAGCTGCTGAACTG
80 90 100 110 120 130 140
GCGTGCACGAGGCGCCAGACCGACGATGATGACCGCCCGCGCGCTGAGGTGCGGTGATGTGAGACAT
140 150 160 170 180 190 200 X
GCTGTCACGAGGCGCCAGACCGACGATGATGACCGCCCGCGCGCTGAGGTGCGGTGATGTGAGACAT
150 160 170 180 190 200 210 X
```

23. US-09-697-123b-12 (1-207)  
US-09-697-123b-7 Sequence 7, Application US/09697123B

Initial Score = 101 Optimized Score = 158 Significance = -0.76  
Residue Identity = 79% Matches = 170 Mismatches = 37  
Gaps = 7 Conservative Substitutions = 0

```
X 10 20 30 40 50 60 70
TCAAGAGAGAGCGCTACGACCTGG-TGCGCTGCGCCGCTTACAAAGTCAACAAAGCTGCTGAACTG
|||||
TCAAGAGAGAGCGCTACGACCTGG-TGCGCTGCGCCGCTTACAAAGTCAACAAAGCTGCTGAACTG
X 10 20 30 40 50 60 70
```

```
GCGAGCC-----GATCACCAGTTGACCGCTGACCGCAAGAGAGATGTCGCCACCATGATACCTGGTTC
|||||
CCACCGCGCTCTGCTGATGCTGACCGACCGCTTACCGAGAGAGAGCTGCTGCCACCATGCTGGTTC
80 90 100 110 120 130 140
GCGTGCACGAGGCGCCAGACCGACGATGATGACCGCCCGCGCGCTGAGGTGCGGTGATGTGAGACAT
140 150 160 170 180 190 200 X
GCTGTCACGAGGCGCCAGACCGACGATGATGACCGCCCGCGCGCTGAGGTGCGGTGATGTGAGACAT
150 160 170 180 190 200 210 X
```

24. US-09-697-123b-12 (1-207)  
US-09-697-123b-1 Sequence 18, Application US/09697123B

Initial Score = 98 Optimized Score = 155 Significance = -0.83  
Residue Identity = 78% Matches = 166 Mismatches = 41  
Gaps = 4 Conservative Substitutions = 0

```
X 10 20 30 40 50 60 70
TCAAGAGAGAGCGCTACGACCTGG-TGCGCTGCGCCGCTTACAAAGTCAACAAAGCTGCTGAACTG
|||||
TCAAGAGAGAGCGCTACGACCTGG-TGCGCTGCGCCGCTTACAAAGTCAACAAAGCTGCTGAACTG
X 10 20 30 40 50 60 70
```



> 0 <  
0110 Intelligenetics  
> 0 <

FastDB - Fast Pairwise Comparison of Sequences  
Release 5.4

Results file us-09-697-123b-11.res made by shanley on Wed 13 Nov 102 14:06:36-PST.

Query sequence being compared: US-09-697-123B-11 (1-208)  
Number of sequences searched: 26  
Number of scores above cutoff: 25

Results of the initial comparison of US-09-697-123B-11 (1-208) with:  
File : US09697123B.seq

```

100-
N -
U -
M -
B -
F -
R -
O -
F 10-
S -
E 5-
Q -
U -
N -
C -
E -
S 0-
SCORE 0 23 46 69 92 116 139 162 185 208
STDDEV -2 -1 0 1 1 1 1 1 1 1

```

## PARAMETERS

Similarity matrix Unitary K-tuple  
Mismatch penalty 1 Joining penalty 4  
Gap penalty 5.00 Window size 30  
Gap size penalty 0.33 207  
Cutoff score 12  
Randomization group 0

## SEARCH STATISTICS

Scores: Mean Median Standard Deviation  
157 181 50.23  
Times: CPU  
00:00:00.00 Total Elapsed  
00:00:00.00

Number of residues: 5077  
Number of sequences searched: 26  
Number of scores above cutoff: 25

The scores below are sorted by initial score.  
Significance is calculated based on initial score.

A 100% identical sequence to the query sequence was found:

Sequence Name	Description	Init. Opt.	Length	Score	Sig.	Frame

1. US-09-697-123B-1 Sequence 11, Application 208 208 208 1.02 0  
The list of other best scores is:

Sequence Name	Description	Length	Score	Init. Opt.	Sig.	Frame
2. US-09-697-123B-1 Sequence 10, Application		208	205	205	0.96	0
3. US-09-697-123B-1 Sequence 15, Application		208	191	191	0.68	0
4. US-09-697-123B-2 Sequence 2, Application U		208	189	189	0.64	0
5. US-09-697-123B-1 Sequence 1, Application U		208	189	189	0.64	0
6. US-09-697-123B-1 Sequence 19, Application		208	187	187	0.60	0
7. US-09-697-123B-1 Sequence 16, Application		208	185	185	0.56	0
8. US-09-697-123B-6 Sequence 6, Application U		208	185	185	0.56	0
9. US-09-697-123B-5 Sequence 5, Application U		208	185	185	0.56	0
10. US-09-697-123B-3 Sequence 3, Application U		208	185	185	0.56	0
11. US-09-697-123B-2 Sequence 22, Application		208	184	184	0.54	0
12. US-09-697-123B-8 Sequence 8, Application U		208	184	184	0.54	0
13. US-09-697-123B-1 Sequence 17, Application		208	181	181	0.48	0
14. US-09-697-123B-2 Sequence 23, Application		208	180	180	0.46	0
15. US-09-697-123B-2 Sequence 24, Application		208	177	177	0.40	0
16. US-09-697-123B-2 Sequence 20, Application		205	166	177	0.18	0
17. US-09-697-123B-1 Sequence 12, Application		207	155	178	-0.04	0
18. US-09-697-123B-4 Sequence 4, Application U		207	151	177	-0.12	0
19. US-09-697-123B-1 Sequence 13, Application		223	141	167	-0.32	0
20. US-09-697-123B-9 Sequence 9, Application U		223	140	168	-0.34	0
21. US-09-697-123B-2 Sequence 21, Application		214	138	159	-0.38	0
22. US-09-697-123B-1 Sequence 14, Application		214	129	166	-0.56	0
23. US-09-697-123B-1 Sequence 18, Application		211	121	166	-0.72	0
24. US-09-697-123B-7 Sequence 7, Application U		214	108	168	-0.98	0
25. US-09-697-123B-2 Sequence 25, Application		19	19	19	-2.75	0

1. US-09-697-123B-11 (1-208)

US-09-697-123B-1 Sequence 11, Application US/09697123B

Initial Score = 208 Optimized Score = 208 Significance = 1.02  
Residue Identity = 100% Matches = 208 Mismatches = 0  
Gaps = 0 Conservative Substitutions = 0

X 10 20 30 40 50 60 70  
TCAAGGAGAGCGCTACGACCTGCGCCGGGTGGCGCGGTACAGGTCAACAGAGACCTGCGCTGAACCGCG  
TCAAGGAGAGCGCTACGACCTGCGCCGGGTGGCGCGGTACAGGTCAACAGAGACCTGCGCTGAACCGCG  
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140  
GCCAGGCCATCACCGCTGACGCTGACCGAGAGAGACGTCGCGCACCATCGATTACCTGCTCGCTTGC  
GCCAGGCCATCACCGCTGACGCTGACCGAGAGAGACGTCGCGCACCATCGATTACCTGCTCGCTTGC  
GCCAGGCCATCACCGCTGACGCTGACCGAGAGAGACGTCGCGCACCATCGATTACCTGCTCGCTTGC  
X 80 90 100 110 120 130 140

150 160 170 180 190 200 X  
ACGAGGGCCAGACCGCGATGACCGCGCGGTGTCGAGGTGTCGAGACGACGAT  
ACGAGGGCCAGACCGCGATGACCGCGCGGTGTCGAGGTGTCGAGACGACGAT  
ACGAGGGCCAGACCGCGATGACCGCGCGGTGTCGAGGTGTCGAGACGACGAT  
X 150 160 170 180 190 200

2. US-09-697-123B-11 (1-208)

US-09-697-123B-1 Sequence 10, Application US/09697123B

Initial Score = 205 Optimized Score = 205 Significance = 0.96  
Residue Identity = 98% Matches = 205 Mismatches = 3  
Gaps = 0 Conservative Substitutions = 0

X 10 20 30 40 50 60 70  
TCAAGGAGAGCGCTACGACCTGCGCCGGGTGGCGCGGTACAGGTCAACAGAGACCTGCGCTGAACCGCG  
TCAAGGAGAGCGCTACGACCTGCGCTGCGGTGCGGTACAGGTCAACAGAGACCTGCGCTGAACCGCG  
X 10 20 30 40 50 60 70



Initial Score = 185 Optimized Score = 185 Significance = 0.56  
Residue Identity = 88% Matches = 185 Mismatches = 23  
Gaps = 0 Conservative Substitutions = 0

X  
TCAGAGAGAGCGCTACGACCTGCGCCCGGCGGCGGCTCAAGGTCAAGTAAGAGCTGCGCTGAAGCGCG  
TCAGAGAGAGAGCGCTACGACCTGCGCCCGGCGGCGGCTCAAGGTCAAGTAAGAGCTGCGCTGAAGCGCG  
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140  
GCCAGCCCATACCACTGACGCTGACCGAGAAAGCTGCTGCGACCATCGAATATCTGCTCCGCTTGC  
GCCAGCCCATACCACTGACGCTGACCGAGAAAGCTGCTGCGACCATCGAATATCTGCTCCGCTTGC  
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140  
GCCAGCCCATACCACTGACGCTGACCGAGAAAGCTGCTGCGACCATCGAATATCTGCTCCGCTTGC  
GCCAGCCCATACCACTGACGCTGACCGAGAAAGCTGCTGCGACCATCGAATATCTGCTCCGCTTGC  
X 10 20 30 40 50 60 70

9. US-09-697-123B-11 (1-208)  
US-09-697-123B-5 Sequence 5, Application US/09697123B

Initial Score = 185 Optimized Score = 185 Significance = 0.56  
Residue Identity = 88% Matches = 185 Mismatches = 23  
Gaps = 0 Conservative Substitutions = 0

X  
TCAGAGAGAGCGCTACGACCTGCGCCCGGCGGCGGCTCAAGGTCAAGTAAGAGCTGCGCTGAAGCGCG  
TCAGAGAGAGAGCGCTACGACCTGCGCCCGGCGGCGGCTCAAGGTCAAGTAAGAGCTGCGCTGAAGCGCG  
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140  
GCCAGCCCATACCACTGACGCTGACCGAGAAAGCTGCTGCGACCATCGAATATCTGCTCCGCTTGC  
GCCAGCCCATACCACTGACGCTGACCGAGAAAGCTGCTGCGACCATCGAATATCTGCTCCGCTTGC  
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140  
GCCAGCCCATACCACTGACGCTGACCGAGAAAGCTGCTGCGACCATCGAATATCTGCTCCGCTTGC  
GCCAGCCCATACCACTGACGCTGACCGAGAAAGCTGCTGCGACCATCGAATATCTGCTCCGCTTGC  
X 10 20 30 40 50 60 70

10. US-09-697-123B-11 (1-208)  
US-09-697-123B-3 Sequence 3, Application US/09697123B

Initial Score = 185 Optimized Score = 185 Significance = 0.56  
Residue Identity = 88% Matches = 185 Mismatches = 23  
Gaps = 0 Conservative Substitutions = 0

X  
TCAGAGAGAGCGCTACGACCTGCGCCCGGCGGCGGCTCAAGGTCAAGTAAGAGCTGCGCTGAAGCGCG  
TCAGAGAGAGAGCGCTACGACCTGCGCCCGGCGGCGGCTCAAGGTCAAGTAAGAGCTGCGCTGAAGCGCG  
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140  
GCCAGCCCATACCACTGACGCTGACCGAGAAAGCTGCTGCGACCATCGAATATCTGCTCCGCTTGC  
GCCAGCCCATACCACTGACGCTGACCGAGAAAGCTGCTGCGACCATCGAATATCTGCTCCGCTTGC  
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140  
GCCAGCCCATACCACTGACGCTGACCGAGAAAGCTGCTGCGACCATCGAATATCTGCTCCGCTTGC  
GCCAGCCCATACCACTGACGCTGACCGAGAAAGCTGCTGCGACCATCGAATATCTGCTCCGCTTGC  
X 10 20 30 40 50 60 70

11. US-09-697-123B-11 (1-208)  
US-09-697-123B-2 Sequence 22, Application US/09697123B

Initial Score = 184 Optimized Score = 184 Significance = 0.54  
Residue Identity = 88% Matches = 184 Mismatches = 24  
Gaps = 0 Conservative Substitutions = 0

X  
TCAGAGAGAGCGCTACGACCTGCGCCCGGCGGCGGCTCAAGGTCAAGTAAGAGCTGCGCTGAAGCGCG  
TCAGAGAGAGAGCGCTACGACCTGCGCCCGGCGGCGGCTCAAGGTCAAGTAAGAGCTGCGCTGAAGCGCG  
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140  
GCCAGCCCATACCACTGACGCTGACCGAGAAAGCTGCTGCGACCATCGAATATCTGCTCCGCTTGC  
GCCAGCCCATACCACTGACGCTGACCGAGAAAGCTGCTGCGACCATCGAATATCTGCTCCGCTTGC  
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140  
GCCAGCCCATACCACTGACGCTGACCGAGAAAGCTGCTGCGACCATCGAATATCTGCTCCGCTTGC  
GCCAGCCCATACCACTGACGCTGACCGAGAAAGCTGCTGCGACCATCGAATATCTGCTCCGCTTGC  
X 10 20 30 40 50 60 70

12. US-09-697-123B-11 (1-208)  
US-09-697-123B-8 Sequence 8, Application US/09697123B

Initial Score = 184 Optimized Score = 184 Significance = 0.54  
Residue Identity = 88% Matches = 184 Mismatches = 24  
Gaps = 0 Conservative Substitutions = 0

X  
TCAGAGAGAGCGCTACGACCTGCGCCCGGCGGCGGCTCAAGGTCAAGTAAGAGCTGCGCTGAAGCGCG  
TCAGAGAGAGAGCGCTACGACCTGCGCCCGGCGGCGGCTCAAGGTCAAGTAAGAGCTGCGCTGAAGCGCG  
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140  
GCCAGCCCATACCACTGACGCTGACCGAGAAAGCTGCTGCGACCATCGAATATCTGCTCCGCTTGC  
GCCAGCCCATACCACTGACGCTGACCGAGAAAGCTGCTGCGACCATCGAATATCTGCTCCGCTTGC  
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140  
GCCAGCCCATACCACTGACGCTGACCGAGAAAGCTGCTGCGACCATCGAATATCTGCTCCGCTTGC  
GCCAGCCCATACCACTGACGCTGACCGAGAAAGCTGCTGCGACCATCGAATATCTGCTCCGCTTGC  
X 10 20 30 40 50 60 70

13. US-09-697-123B-11 (1-208)  
US-09-697-123B-1 Sequence 17, Application US/09697123B

Initial Score = 181 Optimized Score = 181 Significance = 0.48  
Residue Identity = 87% Matches = 181 Mismatches = 27  
Gaps = 0 Conservative Substitutions = 0

X  
TCAGAGAGAGCGCTACGACCTGCGCCCGGCGGCGGCTCAAGGTCAAGTAAGAGCTGCGCTGAAGCGCG  
TCAGAGAGAGAGCGCTACGACCTGCGCCCGGCGGCGGCTCAAGGTCAAGTAAGAGCTGCGCTGAAGCGCG  
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140  
GCCAGCCCATACCACTGACGCTGACCGAGAAAGCTGCTGCGACCATCGAATATCTGCTCCGCTTGC  
GCCAGCCCATACCACTGACGCTGACCGAGAAAGCTGCTGCGACCATCGAATATCTGCTCCGCTTGC  
X 10 20 30 40 50 60 70







```

      80      90      100      110      120      130
GCCAGCC-----CATCACCAGCTGCAGCGTGCAGAGAGAGCGTCGCGCCAGCATCGAATACCTGTC
      |||      |||      |||      |||      |||      |||
CCAACCGGCTGTGTGACTGCCACACGCTCACCAGAGAGACGTCTGCCACCATCGGGTACTGTGTTC
      80      90      100      110      120      130      140
140      150      160      170      180      190      200
GCTTGCAGCAGAGGCCAGACCGCGATGACCGCTCCGGCGGTGTGTGAGGTGGTCGAGACCGAGCAT
|||      |||      |||      |||      |||      |||      |||
GCTTGCAGCAGAGGCCAGACCGCGATGACCGCTCCGGCGGTGTGTGAGGTGGTCGAGCAT
150      160      170      180      190      200      210
X

```

25. US-09-697-123B-11 (1-208)

US-09-697-123B-2 Sequence 25, Application US/09697123B

Initial Score = 19 Optimized Score = 19 Significance = -2.75  
 Residue Identity = 100% Matches = 19 Mismatches = 0  
 Caps = 0 Conservative Substitutions = 0

```

X      10      X
TCAAGGAGAAAGCGCTACGA
|||||      |||||
TCAAGGAGAAAGCGCTACGA
X      10      X

```



80 90 100 110 120 130 140  
GCCAGCCCATTCACGACCTGACGAGTGAAGAGACGCTGCGCCACCTGCAATACCTGTCGCTTGC  
|||||  
GCCAGCCCATTCACGACCTGACGAGTGAAGAGACGCTGCGCCACCTGCAATACCTGTCGCTTGC  
80 90 100 110 120 130 140  
150 160 170 180 190 200 X  
ACGAGGCGCCAGACCGCGATGACCGCTCCGGCGGTGTCGAGTGCCTGACACCGACACAT  
|||||  
ACGAGGCGCCAGACCGCGATGACCGCTCCGGCGGTGTCGAGTGCCTGACACCGACACAT  
150 160 170 180 190 200 X

3. US-09-697-123b-10 (1-208)  
US-09-697-123b-1 Sequence 15, Application US/09697123B

Initial Score = 188 Optimized Score = 188 Significance = 0.64  
Residue Identity = 90% Matches = 188 Mismatches = 20  
Gaps = 0 Conservative Substitutions = 0

X 10 20 30 40 50 60 70  
TCAAGAGAACGGCTACGACCTGCGCTGCGTGCCTGCAAGCTACACAGACCTGCGCTGAACCGCG  
|||||  
TCAAGAGAACGGCTACGACCTGCGCTGCGTGCCTGCAAGCTACACAGACCTGCGCTGAACCGCG  
X 10 20 30 40 50 60 70  
80 90 100 110 120 130 140  
GCCAGCCCATTCACGACCTGACGAGTGAAGAGACGCTGCGCCACCTGCAATACCTGTCGCTTGC  
|||||  
GTGACCGCATACCAAGCTGACGAGTGAAGAGACGCTGCGCCACCTGCAATACCTGTCGCTTGC  
80 90 100 110 120 130 140  
150 160 170 180 190 200 X  
ACGAGGCGCCAGACCGCGATGACCGCTCCGGCGGTGTCGAGTGCCTGACACCGACACAT  
|||||  
ACGAGGCGCCAGACCGCGATGACCGCTCCGGCGGTGTCGAGTGCCTGACACCGACACAT  
150 160 170 180 190 200 X

4. US-09-697-123b-10 (1-208)  
US-09-697-123b-1 Sequence 19, Application US/09697123B

Initial Score = 186 Optimized Score = 186 Significance = 0.60  
Residue Identity = 89% Matches = 186 Mismatches = 22  
Gaps = 0 Conservative Substitutions = 0

X 10 20 30 40 50 60 70  
TCAAGAGAACGGCTACGACCTGCGCTGCGTGCCTGCAAGCTACACAGACCTGCGCTGAACCGCG  
|||||  
TCAAGAGAACGGCTACGACCTGCGCTGCGTGCCTGCAAGCTACACAGACCTGCGCTGAACCGCG  
X 10 20 30 40 50 60 70  
80 90 100 110 120 130 140  
GCCAGCCCATTCACGACCTGACGAGTGAAGAGACGCTGCGCCACCTGCAATACCTGTCGCTTGC  
|||||  
GCCAGCCCATTCACGACCTGACGAGTGAAGAGACGCTGCGCCACCTGCAATACCTGTCGCTTGC  
80 90 100 110 120 130 140  
150 160 170 180 190 200 X  
ACGAGGCGCCAGACCGCGATGACCGCTCCGGCGGTGTCGAGTGCCTGACACCGACACAT  
|||||  
ACGAGGCGCCAGACCGCGATGACCGCTCCGGCGGTGTCGAGTGCCTGACACCGACACAT  
150 160 170 180 190 200 X

5. US-09-697-123b-10 (1-208)  
US-09-697-123b-1 Sequence 16, Application US/09697123B

Initial Score = 186 Optimized Score = 186 Significance = 0.60  
Residue Identity = 89% Matches = 186 Mismatches = 22  
Gaps = 0 Conservative Substitutions = 0

X 10 20 30 40 50 60 70

TCAAGAGAACGGCTACGACCTGCGCTGCGTGCCTGCAAGCTACACAGACCTGCGCTGAACCGCG  
|||||  
TCAAGAGAACGGCTACGACCTGCGCTGCGTGCCTGCAAGCTACACAGACCTGCGCTGAACCGCG  
X 10 20 30 40 50 60 70  
80 90 100 110 120 130 140  
GCCAGCCCATTCACGACCTGACGAGTGAAGAGACGCTGCGCCACCTGCAATACCTGTCGCTTGC  
|||||  
GCCAGCCCATTCACGACCTGACGAGTGAAGAGACGCTGCGCCACCTGCAATACCTGTCGCTTGC  
80 90 100 110 120 130 140  
150 160 170 180 190 200 X  
ACGAGGCGCCAGACCGCGATGACCGCTCCGGCGGTGTCGAGTGCCTGACACCGACACAT  
|||||  
ACGAGGCGCCAGACCGCGATGACCGCTCCGGCGGTGTCGAGTGCCTGACACCGACACAT  
150 160 170 180 190 200 X

6. US-09-697-123b-10 (1-208)  
US-09-697-123b-6 Sequence 6, Application US/09697123B

Initial Score = 186 Optimized Score = 186 Significance = 0.60  
Residue Identity = 89% Matches = 186 Mismatches = 22  
Gaps = 0 Conservative Substitutions = 0

X 10 20 30 40 50 60 70  
TCAAGAGAACGGCTACGACCTGCGCTGCGTGCCTGCAAGCTACACAGACCTGCGCTGAACCGCG  
|||||  
TCAAGAGAACGGCTACGACCTGCGCTGCGTGCCTGCAAGCTACACAGACCTGCGCTGAACCGCG  
X 10 20 30 40 50 60 70  
80 90 100 110 120 130 140  
GCCAGCCCATTCACGACCTGACGAGTGAAGAGACGCTGCGCCACCTGCAATACCTGTCGCTTGC  
|||||  
GCCAGCCCATTCACGACCTGACGAGTGAAGAGACGCTGCGCCACCTGCAATACCTGTCGCTTGC  
80 90 100 110 120 130 140  
150 160 170 180 190 200 X  
ACGAGGCGCCAGACCGCGATGACCGCTCCGGCGGTGTCGAGTGCCTGACACCGACACAT  
|||||  
ACGAGGCGCCAGACCGCGATGACCGCTCCGGCGGTGTCGAGTGCCTGACACCGACACAT  
150 160 170 180 190 200 X

7. US-09-697-123b-10 (1-208)  
US-09-697-123b-5 Sequence 5, Application US/09697123B

Initial Score = 186 Optimized Score = 186 Significance = 0.60  
Residue Identity = 89% Matches = 186 Mismatches = 22  
Gaps = 0 Conservative Substitutions = 0

X 10 20 30 40 50 60 70  
TCAAGAGAACGGCTACGACCTGCGCTGCGTGCCTGCAAGCTACACAGACCTGCGCTGAACCGCG  
|||||  
TCAAGAGAACGGCTACGACCTGCGCTGCGTGCCTGCAAGCTACACAGACCTGCGCTGAACCGCG  
X 10 20 30 40 50 60 70  
80 90 100 110 120 130 140  
GCCAGCCCATTCACGACCTGACGAGTGAAGAGACGCTGCGCCACCTGCAATACCTGTCGCTTGC  
|||||  
GCCAGCCCATTCACGACCTGACGAGTGAAGAGACGCTGCGCCACCTGCAATACCTGTCGCTTGC  
80 90 100 110 120 130 140  
150 160 170 180 190 200 X  
ACGAGGCGCCAGACCGCGATGACCGCTCCGGCGGTGTCGAGTGCCTGACACCGACACAT  
|||||  
ACGAGGCGCCAGACCGCGATGACCGCTCCGGCGGTGTCGAGTGCCTGACACCGACACAT  
150 160 170 180 190 200 X

8. US-09-697-123b-10 (1-208)  
US-09-697-123b-2 Sequence 2, Application US/09697123B

Initial Score = 186 Optimized Score = 186 Significance = 0.60  
Residue Identity = 89% Matches = 186 Mismatches = 22  
Gaps = 0 Conservative Substitutions = 0

X  
TCAGAGGAAGCCCTACGACCTGCTGCGTGGTGTGCTGACAGGTCAACAGAGAGCTGCGCTGAACGCCG  
|||||  
TCAGAGGAAGCCCTACGACCTGCTGCGTGGTGTGCTGACAGGTCAACAGAGAGCTGCGCTGAACGCCG  
X  
10 20 30 40 50 60 70  
10 20 30 40 50 60 70

80 90 100 110 120 130 140  
GCCAGCCCATCACCACCTGACGCTGACCGAGAGAGAGCTGCTGCGACCATCAATACGCTGCTGCGCTGCG  
|||||  
GCCAGCCCATCACCACCTGACGCTGACCGAGAGAGAGCTGCTGCGACCATCAATACGCTGCTGCGCTGCG  
X  
80 90 100 110 120 130 140

150 160 170 180 190 200  
ACGAGGGCCGACGACGCTGACCGGCGGCGGTGTGAGAGTGGCGGTGACAGACGACAT X  
|||||  
ACGAGGCTAGTGGCGATGACGCTGCGGCGGCGGTGTGAGAGTGGCGGTGACAGACGACAT X  
150 160 170 180 190 200

## 9. US-09-697-123B-10 (1-208)

US-09-697-123B-1 Sequence 1, Application US/09697123B

Initial Score = 186 Optimized Score = 186 Significance = 0.60  
Residue Identity = 89% Matches = 186 Mismatches = 22  
Gaps = 0 Conservative Substitutions = 0

X  
TCAGAGGAAGCCCTACGACCTGCTGCGTGGTGTGCTGACAGGTCAACAGAGAGCTGCGCTGAACGCCG  
|||||  
TCAGAGGAAGCCCTACGACCTGCTGCGTGGTGTGCTGACAGGTCAACAGAGAGCTGCGCTGAACGCCG  
X  
10 20 30 40 50 60 70  
10 20 30 40 50 60 70

80 90 100 110 120 130 140  
GCCAGCCCATCACCACCTGACGCTGACCGAGAGAGAGCTGCTGCGACCATCAATACGCTGCTGCGCTGCG  
|||||  
GCCAGCCCATCACCACCTGACGCTGACCGAGAGAGAGCTGCTGCGACCATCAATACGCTGCTGCGCTGCG  
X  
80 90 100 110 120 130 140

150 160 170 180 190 200  
ACGAGGGCCGACGACGCTGACCGGCGGCGGTGTGAGAGTGGCGGTGACAGACGACAT X  
|||||  
ACGAGGCTAGTGGCGATGACGCTGCGGCGGCGGTGTGAGAGTGGCGGTGACAGACGACAT X  
150 160 170 180 190 200

## 10. US-09-697-123B-10 (1-208)

US-09-697-123B-2 Sequence 22, Application US/09697123B

Initial Score = 185 Optimized Score = 185 Significance = 0.58  
Residue Identity = 88% Matches = 185 Mismatches = 23  
Gaps = 0 Conservative Substitutions = 0

X  
TCAGAGGAAGCCCTACGACCTGCTGCGTGGTGTGCTGACAGGTCAACAGAGAGCTGCGCTGAACGCCG  
|||||  
TCAGAGGAAGCCCTACGACCTGCTGCGTGGTGTGCTGACAGGTCAACAGAGAGCTGCGCTGAACGCCG  
X  
10 20 30 40 50 60 70  
10 20 30 40 50 60 70

80 90 100 110 120 130 140  
GCCAGCCCATCACCACCTGACGCTGACCGAGAGAGAGCTGCTGCGACCATCAATACGCTGCTGCGCTGCG  
|||||  
GCCAGCCCATCACCACCTGACGCTGACCGAGAGAGAGCTGCTGCGACCATCAATACGCTGCTGCGCTGCG  
X  
80 90 100 110 120 130 140

150 160 170 180 190 200  
ACGAGGGCCGACGACGCTGACCGGCGGCGGTGTGAGAGTGGCGGTGACAGACGACAT X  
|||||  
ACGAGGCTAGTGGCGATGACGCTGCGGCGGCGGTGTGAGAGTGGCGGTGACAGACGACAT X  
150 160 170 180 190 200

## 11. US-09-697-123B-10 (1-208)

US-09-697-123B-3 Sequence 3, Application US/09697123B

Initial Score = 183 Optimized Score = 183 Significance = 0.54  
Residue Identity = 87% Matches = 183 Mismatches = 25  
Gaps = 0 Conservative Substitutions = 0

X  
TCAGAGGAAGCCCTACGACCTGCTGCGTGGTGTGCTGACAGGTCAACAGAGAGCTGCGCTGAACGCCG  
|||||  
TCAGAGGAAGCCCTACGACCTGCTGCGTGGTGTGCTGACAGGTCAACAGAGAGCTGCGCTGAACGCCG  
X  
10 20 30 40 50 60 70  
10 20 30 40 50 60 70

80 90 100 110 120 130 140  
GCCAGCCCATCACCACCTGACGCTGACCGAGAGAGAGCTGCTGCGACCATCAATACGCTGCTGCGCTGCG  
|||||  
GCCATCCGATCAACACCTGACGCTGACCGAGAGAGAGCTGCTGCGACCATCAATACGCTGCTGCGCTGCG  
X  
80 90 100 110 120 130 140

150 160 170 180 190 200  
ACGAGGGCCGACGACGCTGACCGGCGGCGGTGTGAGAGTGGCGGTGACAGACGACAT X  
|||||  
ACGAGGCTAGTGGCGATGACGCTGCGGCGGCGGTGTGAGAGTGGCGGTGACAGACGACAT X  
150 160 170 180 190 200

## 12. US-09-697-123B-10 (1-208)

US-09-697-123B-8 Sequence 8, Application US/09697123B

Initial Score = 182 Optimized Score = 182 Significance = 0.52  
Residue Identity = 87% Matches = 182 Mismatches = 26  
Gaps = 0 Conservative Substitutions = 0

X  
TCAGAGGAAGCCCTACGACCTGCTGCGTGGTGTGCTGACAGGTCAACAGAGAGCTGCGCTGAACGCCG  
|||||  
TCAGAGGAAGCCCTACGACCTGCTGCGTGGTGTGCTGACAGGTCAACAGAGAGCTGCGCTGAACGCCG  
X  
10 20 30 40 50 60 70  
10 20 30 40 50 60 70

80 90 100 110 120 130 140  
GCCAGCCCATCACCACCTGACGCTGACCGAGAGAGAGCTGCTGCGACCATCAATACGCTGCTGCGCTGCG  
|||||  
ATCATCCGATCAACACCTGACGCTGACCGAGAGAGAGCTGCTGCGACCATCAATACGCTGCTGCGCTGCG  
X  
80 90 100 110 120 130 140

150 160 170 180 190 200  
ACGAGGGCCGACGACGCTGACCGGCGGCGGTGTGAGAGTGGCGGTGACAGACGACAT X  
|||||  
ACGAGGCTAGTGGCGATGACGCTGCGGCGGCGGTGTGAGAGTGGCGGTGACAGACGACAT X  
150 160 170 180 190 200

## 13. US-09-697-123B-10 (1-208)

US-09-697-123B-2 Sequence 23, Application US/09697123B

Initial Score = 179 Optimized Score = 179 Significance = 0.46  
Residue Identity = 86% Matches = 179 Mismatches = 29  
Gaps = 0 Conservative Substitutions = 0

X  
TCAGAGGAAGCCCTACGACCTGCTGCGTGGTGTGCTGACAGGTCAACAGAGAGCTGCGCTGAACGCCG  
|||||  
TCAGAGGAAGCCCTACGACCTGCTGCGTGGTGTGCTGACAGGTCAACAGAGAGCTGCGCTGAACGCCG  
X  
10 20 30 40 50 60 70  
10 20 30 40 50 60 70

80 90 100 110 120 130 140  
GCCAGCCCATCACCACCTGACGCTGACCGAGAGAGAGCTGCTGCGACCATCAATACGCTGCTGCGCTGCG  
|||||  
GTGAGCCGATCAACACCTGACGCTGACCGAGAGAGAGCTGCTGCGACCATCAATACGCTGCTGCGCTGCG  
X  
80 90 100 110 120 130 140

X 10 20 30 40 50 60 70  
TCAAGAGAAAGCGCTACGACCTGGCTGGCTGGCTACAGGTCAACAGAAGCTGGGCTGAACGCCG

```
|||||
TCAAGGAGAAAGCGTACGACCTGGCCCGCTGGCCGCTACAGGTCAACAGAAAGCTGGCGCTGACACCG
X 10 20 30 40 50 60 70
80 90 100 110 120 130 140
GCCACCCCATTCACAGCTTCGACGCTGACGAGAGAGACGTCTGCGCCACCAATTCGATCTCCGCTTGC
|||
ATCATTCGCAATCACACACGACGAGCTGACGAGAAAGACGTCTGCGCCACCAATTCGATCTCCGCTTGC
80 90 100 110 120 130 140
150 160 170 180 190 200 210
ACGAGCGC-----CAGACCGCGATGACCGCTCGCGCGCTGTCGAGTCCGCTGACAGCG
|||
ACGAGCGCTTCAGAGGTGGCCAGCGCCCGCTTATGACTCTCCCGCGGGGTCCAGAGTCCCGGTGAACCG
150 160 170 180 190 200 210
X
ACGACAT
|||||
ACGACAT
220 X
```

20. US-09-697-123B-10 (1-208)  
US-09-697-123B-2 Sequence 21, Application US/09697123B

Initial Score = 139 Optimized Score = 170 Significance = -0.34  
Residue Identity = 82% Matches = 177 Mismatches = 31  
Gaps = 6 Conservative Substitutions = 0

```
X 10 20 30 40 50 60 70
TCAAGGAGAAAGCGCTACGACCTGGCTGCGGTGCTGCTACAGGTCAACAGAAAGCTCCGCTGACGCG
|||||
TCAAGGAGAAAGCGCTACGACCTGGCTGCGGTGCTGCTACAGGTCAACAGAAAGCTCCGCTGACGCG
X 10 20 30 40 50 60 70
80 90 100 110 120 130 140
GCCAGC-----CATCACCAGCTTCGACGCTGACCGAGAGAGAGCTGTCGCCACCTACGATGCTGTC
|||||
CCAATTCGCGCTTCAGAGTACACACACCTTCACAGGAGAAAGCTGTCGCCACCTACGATGCTGTC
80 90 100 110 120 130 140
```

```
140 150 160 170 180 190 200
GCTTTCACAGAGGCGACCGCGATGACCGCTCGGGGCTGTCGAGGTCCGCTGAGACCGACGACAT
|||||
GCCGTGCAGAGGCGACCGCGATGACCGCGCTCGGGGCTGTCGAGGTCCGCTGAGATGTCGATGACGACAT
150 160 170 180 190 200 210 X
```

21. US-09-697-123B-10 (1-208)  
US-09-697-123B-9 Sequence 9, Application US/09697123B

Initial Score = 139 Optimized Score = 167 Significance = -0.34  
Residue Identity = 79% Matches = 177 Mismatches = 31  
Gaps = 15 Conservative Substitutions = 0

```
X 10 20 30 40 50 60 70
TCAAGGAGAAAGCGCTACGACCTGGCTGCGGTGCTGCTACAGGTCAACAGAAAGCTCCGCTGACGCG
|||||
TCAAGGAGAAAGCGCTACGACCTGGCTGCGGTGCTGCTACAGGTCAACAGAAAGCTCCGCTGACGCG
X 10 20 30 40 50 60 70
80 90 100 110 120 130 140
GCCAGCCCATTCACAGCTTCGACGCTGACGAGAGAGAGCTGTCGTCGACCATTCGATCTCCGCTTGC
|||||
GGAGCGCATTCAGCTTCGACGCTGACGAGAGAGAGCTGTCGTCGACCATTCGATCTCCGCTTGC
80 90 100 110 120 130 140
```

```
150 160 170 180 190 200
ACGAGCGCGAGA-----CCGCGATGACCGCTCGGGGCTGTCGAGGTCCGCTGAGACCG
|||||
ACGAGCGCGATTCAGGTGCGACCGCGCTGATCTCCCGCGGCGCATTCGAGGTCCGCTGAGACCG
150 160 170 180 190 200 210
```

```
X
ACGACAT
|||||
ACGACAT
220 X
```

22. US-09-697-123B-10 (1-208)  
US-09-697-123B-1 Sequence 14, Application US/09697123B

Initial Score = 126 Optimized Score = 163 Significance = -0.60  
Residue Identity = 79% Matches = 170 Mismatches = 38  
Gaps = 6 Conservative Substitutions = 0

```
X 10 20 30 40 50 60 70
TCAAGGAGAAAGCGCTACGACCTGGCTGCGGTGCTGCTACAGGTCAACAGAAAGCTCCGCTGACGCG
|||||
TCAAGGAGAAAGCGCTACGACCTGGCTGCGGTGCTGCTACAGGTCAACAGAAAGCTCCGCTGACGCG
X 10 20 30 40 50 60 70
```

```
80 90 100 110 120 130
GCCAGC-----CATCACCAGCTTCGACGCTGACCGAGAGAGAGCTGTCGCCACCTACGATGCTGTC
|||||
CCGATTCGCGCTTCAGAGTACACACCTTCACAGGAGAAAGCTGTCGCCACCTACGATGCTGTC
80 90 100 110 120 130 140
```

```
140 150 160 170 180 190 200
GCTTTCACAGAGGCGACCGCGATGACCGCTCGGGGCTGTCGAGGTCCGCTGAGACCGACGACAT
|||||
GCTTTCACAGAGGCGACCGCGATGACCGCTCGGGGCTGTCGAGGTCCGCTGAGACCGACGACAT
150 160 170 180 190 200 210 X
```

23. US-09-697-123B-10 (1-208)  
US-09-697-123B-1 Sequence 18, Application US/09697123B

Initial Score = 122 Optimized Score = 167 Significance = -0.68  
Residue Identity = 81% Matches = 173 Mismatches = 35  
Gaps = 3 Conservative Substitutions = 0

```
X 10 20 30 40 50 60 70
TCAAGGAGAAAGCGCTACGACCTGGCTGCGGTGCTGCTACAGGTCAACAGAAAGCTCCGCTGACGCG
|||||
TCAAGGAGAAAGCGCTACGACCTGGCTGCGGTGCTGCTACAGGTCAACAGAAAGCTCCGCTGACGCG
X 10 20 30 40 50 60 70
80 90 100 110 120 130 140
GCCAGCCCATTCACAGCTTCGACGCTGACGAGAGAGAGCTGTCGTCGACCATTCGATCTCCGCTTGC
|||||
ACCGCGCGACACGACCTTCGACCGAGAGAGAGAGCTGTCGTCGACCATTCGATCTCCGCTTGC
80 90 100 110 120 130 140
```

```
150 160 170 180 190 200
TTCACGAGGCGCGACGATGACCGCTCGGGGCTGTCGAGGTGCTGTCGACCGACGACAT
|||||
TTCACGAGGCGCGACGATGACCGCTCGGGGCTGTCGAGGTGCTGTCGACCGACGACAT
150 160 170 180 190 200 210
```

24. US-09-697-123B-10 (1-208)  
US-09-697-123B-7 Sequence 7, Application US/09697123B

Initial Score = 107 Optimized Score = 167 Significance = -0.98  
Residue Identity = 81% Matches = 174 Mismatches = 34  
Gaps = 6 Conservative Substitutions = 0

```
X 10 20 30 40 50 60 70
TCAAGGAGAAAGCGCTACGACCTGGCTGCGGTGCTGCTACAGGTCAACAGAAAGCTCCGCTGACGCG
|||||
TCAAGGAGAAAGCGCTACGACCTGGCTGCGGTGCTGCTACAGGTCAACAGAAAGCTCCGCTGACGCG
X 10 20 30 40 50 60 70
```



25. US-09-697-123B-10 (1-208)  
US-09-697-123B-2 Sequence 25, Application US/09697123B

Initial Score	-	19	Optimized Score	-	19	Significance	-	-2.75
Residue Identity	-	100%	Matches	-	19	Mismatches	-	0
Gaps	-	0	Conservative Substitutions	-			-	0

  

X	10	X
TCACGAGAGAGCGCTACGA		
TCACGAGAGAGCGCTACGA		
X	10	X

FastDB - Fast Pairwise Comparison of Sequences  
Release 5.4

Query sequence being compared:	US-09-657-123B-9 (1-1223)
Number of sequences searched:	26
Number of scores above cutoff:	25

Results of the initial comparison of US-09-697-123B-9 (1-223) with  
File : US09697123B.seq

Letter	Score
N	-1
U	125
M	50
B	-1
E	74
R	99
O	0
F	124
S	149
E	173
O	198
U	223
E	1
N	1
C	1
E	1
S	1

	unifary	k-tuple
Similarity matrix	1	4
Mismatch penalty	5.00	30
Gap penalty	0.33	Window size
Gap size penalty	13	207
Cutoff score	0	
Randomization group		

Scores:	Mean	Median	Standard Deviation
	1.07	1.00	1.12

Times:	CPU	Total Elapsed
0.000000	0.000000	0.000000

00:00:00.00

Number of residues: 5077

Number of sequences searched:  
Number of scores above cutoff:

The scores below are sorted by initial score.  
Significance is calculated based on initial score

A 100% identical sequence to the query sequence was found:

Sequence Name	Description	Length	Init. Score	Opt. Score	Sig.	Frame
---------------	-------------	--------	-------------	------------	------	-------

1. US-09-697-123B-9 Sequence 9, Application U	223	223	2.17	0
---	-----	-----	------	---

Sequence Name	Description	Length	Init. Opt. Score	Sig. Frame
---------------	-------------	--------	------------------	------------

[illegible]

1. US-09-697-123B-9 (1-223)  
US-09-697-123B-9 Sequencia 9  
US-09-697-123B-9 Sequencia 9

Initial Score	-	223	Optimized Score	-	223	Significance	-	2.17
Residue Identity	-	100%	Matches	-	223	Mismatches	-	0
Gaps	-	0	Conservative Substitutions	-			-	0

TC AAGAGAGGCGCTACACAGCTGGCCCGCGCTGGCCCGCTACAAAGTCAACAAAGAGTGGTTCGACCCCGC  
10 20 30 40 50 60 70  
TC AAGAGAGGCGCTACACAGCTGGCCCGCGCTGGCCCGCTACAAAGTCAACAAAGTGGTTCGACCCCGC  
10 20 30 40 50 60 70  
TC AAGAGAGGCGCTACACAGCTGGCCCGCGCTGGCCCGCTACAAAGTCAACAAAGTGGTTCGACCCCGC  
10 20 30 40 50 60 70

GCAGCGCATCAGTCGTCGCACGGTACCGAGAAAGACGTCGCGACATCGAATACCTGTTCCGGCTGG  
80 90 100 110 120 130 140  
GCAGCGCATCAGTCGTCGCACGGTACCGAGAAAGACGTCGCGACATCGAATACCTGTTCCGGCTGG  
80 90 100 110 120 130 140

ACCAGCCCGTATGGCAGCCGCGCTATGCTCCCGCGCATGAGTGCGGTGAGACCC  
150 160 170 180 190 200 210  
ACAGCCCGTATGGATGGCAGCCCGCTATGACTCTCCGCGCATGAGTGCGGTGAGACCC  
150 160 170 180 190 200 210

220 X  
ACGACAT  
111111  
ACGACAT  
220 X

2. US-09-697-123B-9 (1-223)

US-09-697-123B-1 Sequence 13, Application US/09697123B

```
Initial Score - 195 Optimized Score - 195 Significance - 1.54
```

X 10 20 30 40 50 60 70  
 TCTAAGGAGAGCGGTACGACCTGGCCCGCTGCAAGGCAACAGAAAGCTGGTGTGCAGCGCC  
 X 10 20 30 40 50 60 70  
 TCTAAGGAGAGCGGTACGACCTGGCCCGCTGCAAGGCAACAGAAAGCTGGTGTGCAGCGCC

ACCAAGCCCCATGCGATGGCCACGCCCGGTACATGACTCTCCCGGGCGCATGAGAGTCCCGGTGAGACC  
150 160 170 180 190 200 210

ACGAGGCTCAGA-----CCACGATGACCGTTCGGGGCGCTCGAGGTGCGCGGTGGAAACCG  
150 160 170 180 190 200  
220 X  
ACGACAT  
|||||  
ACGACAT  
X

7. US-09-697-123B-9 (1-223)  
US-09-697-123B-6 Sequence 6, Application US/09697123B

Initial Score = 144 Optimized Score = 170 Significance = 0.39  
Residue Identity = 80% Matches = 180 Mismatches = 28  
Gaps = 15 Conservative Substitutions = 0

X 10 20 30 40 50 60 70  
TCAAGAGAAAGCGCTACGACCTGCGCCCGCGCTGCGCGCTACAAAGGTCAAAAGAGTGTGTCGACGCG  
|||||  
TCAAGAGAAAGCGCTACGACCTGCGCCCGCGCTGCGCGCTATTAAGGTCAAAAGAGTGTGTCGACGCG  
X 10 20 30 40 50 60 70  
GCGAGCGGATCAGCTGTGTCACGCTGACGAGAGAGAGTGTGCGCGCATTCGAATACCTGTGTCGCGCTGC  
|||||  
GCGAGCGGATCAGCTGTGTCACGCTGACGAGAGAGAGTGTGCGCGCATTCGAATACCTGTGTCGCGCTGC  
80 90 100 110 120 130 140  
ACCAAGCCCGCTACGAGATGGCGAGCCCGCGCTACGTGTCGCGCGCATTCGAATACCTGTGTCGAGACCG  
|||||  
ACGAGGCTCAGA-----CCACGATGACCGTTCGGGGCGCTCGAGGTGCGCGGTGGAAACCG  
150 160 170 180 190 200 210  
220 X  
ACGACAT  
|||||  
ACGACAT  
X

8. US-09-697-123B-9 (1-223)  
US-09-697-123B-5 Sequence 5, Application US/09697123B

Initial Score = 144 Optimized Score = 170 Significance = 0.39  
Residue Identity = 80% Matches = 180 Mismatches = 28  
Gaps = 15 Conservative Substitutions = 0

X 10 20 30 40 50 60 70  
TCAAGAGAAAGCGCTACGACCTGCGCCCGCGCTGCGCGCTACAAAGGTCAAAAGAGTGTGTCGACGCG  
|||||  
TCAAGAGAAAGCGCTACGACCTGCGCCCGCGCTGCGCGCTATTAAGGTCAAAAGAGTGTGTCGACGCG  
X 10 20 30 40 50 60 70  
GCGAGCGGATCAGCTGTGTCACGCTGACGAGAGAGAGTGTGCGCGCATTCGAATACCTGTGTCGCGCTGC  
|||||  
GCGAGCGGATCAGCTGTGTCACGCTGACGAGAGAGAGTGTGCGCGCATTCGAATACCTGTGTCGCGCTGC  
80 90 100 110 120 130 140  
ACCAAGCCCGCTACGAGATGGCGAGCCCGCGCTACGTGTCGCGCGCATTCGAATACCTGTGTCGAGACCG  
|||||  
ACGAGGCTCAGA-----CCACGATGACCGTTCGGGGCGCTCGAGGTGCGCGGTGGAAACCG  
150 160 170 180 190 200 210  
220 X  
ACGACAT  
|||||  
ACGACAT  
X

9. US-09-697-123B-9 (1-223)  
US-09-697-123B-2 Sequence 2, Application US/09697123B

Initial Score = 144 Optimized Score = 170 Significance = 0.39  
Residue Identity = 80% Matches = 180 Mismatches = 28  
Gaps = 15 Conservative Substitutions = 0

X 10 20 30 40 50 60 70  
TCAAGAGAAAGCGCTACGACCTGCGCCCGCGCTGCGCGCTACAAAGGTCAAAAGAGTGTGTCGACGCG  
|||||  
TCAAGAGAAAGCGCTACGACCTGCGCCCGCGCTGCGCGCTATTAAGGTCAAAAGAGTGTGTCGACGCG  
X 10 20 30 40 50 60 70  
GCGAGCGGATCAGCTGTGTCACGCTGACGAGAGAGAGTGTGCGCGCATTCGAATACCTGTGTCGCGCTGC  
|||||  
GCGAGCGGATCAGCTGTGTCACGCTGACGAGAGAGAGTGTGCGCGCATTCGAATACCTGTGTCGCGCTGC  
80 90 100 110 120 130 140  
ACCAAGCCCGCTACGAGATGGCGAGCCCGCGCTACGTGTCGCGCGCATTCGAATACCTGTGTCGAGACCG  
|||||  
ACGAGGCTCAGTGC-----GGGATGACCGTTCGCGCGCATTCGAATACCTGTGTCGAGACCG  
150 160 170 180 190 200  
220 X  
ACGACAT  
|||||  
ACGACAT  
X

10. US-09-697-123B-9 (1-223)  
US-09-697-123B-2 Sequence 22, Application US/09697123B

Initial Score = 143 Optimized Score = 169 Significance = 0.36  
Residue Identity = 80% Matches = 179 Mismatches = 29  
Gaps = 15 Conservative Substitutions = 0

X 10 20 30 40 50 60 70  
TCAAGAGAAAGCGCTACGACCTGCGCCCGCGCTGCGCGCTACAAAGGTCAAAAGAGTGTGTCGACGCG  
|||||  
TCAAGAGAAAGCGCTACGACCTGCGCCCGCGCTGCGCGCTATTAAGGTCAAAAGAGTGTGTCGACGCG  
X 10 20 30 40 50 60 70  
GCGAGCGGATCAGCTGTGTCACGCTGACGAGAGAGAGTGTGCGCGCATTCGAATACCTGTGTCGCGCTGC  
|||||  
GCGAGCGGATCAGCTGTGTCACGCTGACGAGAGAGAGTGTGCGCGCATTCGAATACCTGTGTCGCGCTGC  
80 90 100 110 120 130 140  
ACCAAGCCCGCTACGAGATGGCGAGCCCGCGCTACGTGTCGCGCGCATTCGAATACCTGTGTCGAGACCG  
|||||  
ACGAGGCTCAGA-----CCACGATGATCGTTCGGGGCGCGGTGAGGTGCGCGGTGGAAACCG  
150 160 170 180 190 200  
220 X  
ACGACAT  
|||||  
ACGACAT  
X

11. US-09-697-123B-9 (1-223)  
US-09-697-123B-1 Sequence 11, Application US/09697123B

Initial Score = 140 Optimized Score = 168 Significance = 0.29  
Residue Identity = 79% Matches = 178 Mismatches = 30  
Gaps = 15 Conservative Substitutions = 0

X 10 20 30 40 50 60 70

TCAGAGAGAGCCCTACGACCTGCGCCGCGCTACAGGTCAACAGAGCTGGTCTGCACGCCG  
|||||  
TCAGAGAGAGCCCTACGACCTGCGCCGCGCTACAGGTCAACAGAGCTGGTCTGCACGCCG  
X 10 20 30 40 50 60 70  
GCGAGCCGATCAGCTGCTCCACGCTACGAGAGAGAGCTGCGGACCATGAAATACCTGGTGGCTGC  
|||||  
GCGAGCCGATCAGCTGCTCCACGCTACGAGAGAGAGCTGCGGACCATGAAATACCTGGTGGCTGC  
80 90 100 110 120 130 140  
ACCAAGCCGCTACGAGATGGCCAGCCCGCTCATGACTGTCCCGGCGCATCGAGAGCTCGGTGAGACCG  
|||||  
ACGAGGCGCCAGA-----CCGCGATGACCCGCTCGGCGGTGTCAGAGTGGCGGTGAGACCG  
150 160 170 180 190 200  
220 X  
ACGACAT  
|||||  
ACGACAT  
X  
12. US-09-697-123B-9 (1-223)  
US-09-697-123B-1 Sequence 10, Application US/09697123B  
Initial Score = 139 Optimized Score = 167 Significance = 0.27  
Residue Identity = 79% Matches = 177 Mismatches = 31  
Gaps = 15 Conservative Substitutions = 0  
X 10 20 30 40 50 60 70  
TCAGAGAGAGCCGCTACGACCTGCGCCGCGCTACAGGTCAACAGAGCTGGTCTGCACGCCG  
|||||  
TCAGAGAGAGCCGCTACGACCTGCGCCGCGCTACAGGTCAACAGAGCTGGTCTGCACGCCG  
X 10 20 30 40 50 60 70  
TCAGAGAGAGCCGCTACGACCTGCGCCGCGCTACAGGTCAACAGAGCTGGTCTGCACGCCG  
80 90 100 110 120 130 140  
GCGAGCCGATCAGCTGCTCCACGCTACGAGAGAGAGCTGCGGACCATGAAATACCTGGTGGCTGC  
|||||  
GCGAGCCGATCAGCTGCTCCACGCTACGAGAGAGAGCTGCGGACCATGAAATACCTGGTGGCTGC  
150 160 170 180 190 200  
ACCAAGCCGCTACGAGATGGCCAGCCCGCTCATGACTGTCCCGGCGCATCGAGAGCTCGGTGAGACCG  
|||||  
ACGAGGCGCCAGA-----CCGCGATGACCCGCTCGGCGGTGTCAGAGTGGCGGTGAGACCG  
150 160 170 180 190 200  
220 X  
ACGACAT  
|||||  
ACGACAT  
X  
13. US-09-697-123B-9 (1-223)  
US-09-697-123B-2 Sequence 23, Application US/09697123B  
Initial Score = 138 Optimized Score = 163 Significance = 0.25  
Residue Identity = 77% Matches = 173 Mismatches = 35  
Gaps = 15 Conservative Substitutions = 0  
X 10 20 30 40 50 60 70  
TCAGAGAGAGCCGCTACGACCTGCGCCGCGCTACAGGTCAACAGAGCTGGTCTGCACGCCG  
|||||  
TCAGAGAGAGCCGCTACGACCTGCGCCGCGCTACAGGTCAACAGAGCTGGTCTGCACGCCG  
X 10 20 30 40 50 60 70  
TCAGAGAGAGCCGCTACGACCTGCGCCGCGCTACAGGTCAACAGAGCTGGTCTGCACGCCG  
80 90 100 110 120 130 140  
GCGAGCCGATCAGCTGCTCCACGCTACGAGAGAGAGCTGCGGACCATGAAATACCTGGTGGCTGC  
|||||  
GCGAGCCGATCAGCTGCTCCACGCTACGAGAGAGAGCTGCGGACCATGAAATACCTGGTGGCTGC  
150 160 170 180 190 200  
ACCAAGCCGCTACGAGATGGCCAGCCCGCTCATGACTGTCCCGGCGCATCGAGAGCTCGGTGAGACCG  
|||||  
ACGAGGCGCCAGA-----CCGCGATGACCCGCTCGGCGGTGTCAGAGTGGCGGTGAGACCG  
150 160 170 180 190 200  
220 X  
ACGACAT  
|||||  
ACGACAT  
X

80 90 100 110 120 130 140  
ACCAAGCCGCTACGAGATGGCCAGCCCGCTCATGACTGTCCCGGCGCATCGAGAGCTCGGTGAGACCG  
|||||  
ACGAGGCGCCAGA-----CCGCGATGACCCGCTCGGCGGTGTCAGAGTGGCGGTGAGACCG  
150 160 170 180 190 200  
220 X  
ACGACAT  
|||||  
ACGACAT  
X  
14. US-09-697-123B-8 (1-223)  
US-09-697-123B-8 Sequence 8, Application US/09697123B  
Initial Score = 138 Optimized Score = 165 Significance = 0.25  
Residue Identity = 78% Matches = 175 Mismatches = 33  
Gaps = 15 Conservative Substitutions = 0  
X 10 20 30 40 50 60 70  
TCAGAGAGAGCCGCTACGACCTGCGCCGCGCTACAGGTCAACAGAGCTGGTCTGCACGCCG  
|||||  
TCAGAGAGAGCCGCTACGACCTGCGCCGCGCTACAGGTCAACAGAGCTGGTCTGCACGCCG  
X 10 20 30 40 50 60 70  
TCAGAGAGAGCCGCTACGACCTGCGCCGCGCTACAGGTCAACAGAGCTGGTCTGCACGCCG  
80 90 100 110 120 130 140  
GCGAGCCGATCAGCTGCTCCACGCTACGAGAGAGAGCTGCGGACCATGAAATACCTGGTGGCTGC  
|||||  
GCGAGCCGATCAGCTGCTCCACGCTACGAGAGAGAGCTGCGGACCATGAAATACCTGGTGGCTGC  
150 160 170 180 190 200  
ACCAAGCCGCTACGAGATGGCCAGCCCGCTCATGACTGTCCCGGCGCATCGAGAGCTCGGTGAGACCG  
|||||  
ACGAGGCGCCAGA-----CCGCGATGACCCGCTCGGCGGTGTCAGAGTGGCGGTGAGACCG  
150 160 170 180 190 200  
220 X  
ACGACAT  
|||||  
ACGACAT  
X  
15. US-09-697-123B-9 (1-223)  
US-09-697-123B-1 Sequence 17, Application US/09697123B  
Initial Score = 137 Optimized Score = 164 Significance = 0.23  
Residue Identity = 78% Matches = 174 Mismatches = 34  
Gaps = 15 Conservative Substitutions = 0  
X 10 20 30 40 50 60 70  
TCAGAGAGAGCCGCTACGACCTGCGCCGCGCTACAGGTCAACAGAGCTGGTCTGCACGCCG  
|||||  
TCAGAGAGAGCCGCTACGACCTGCGCCGCGCTACAGGTCAACAGAGCTGGTCTGCACGCCG  
X 10 20 30 40 50 60 70  
TCAGAGAGAGCCGCTACGACCTGCGCCGCGCTACAGGTCAACAGAGCTGGTCTGCACGCCG  
80 90 100 110 120 130 140  
GCGAGCCGATCAGCTGCTCCACGCTACGAGAGAGAGCTGCGGACCATGAAATACCTGGTGGCTGC  
|||||  
GCGAGCCGATCAGCTGCTCCACGCTACGAGAGAGAGCTGCGGACCATGAAATACCTGGTGGCTGC  
150 160 170 180 190 200  
ACCAAGCCGCTACGAGATGGCCAGCCCGCTCATGACTGTCCCGGCGCATCGAGAGCTCGGTGAGACCG  
|||||  
ACGAGGCGCCAGA-----CCGCGATGACCCGCTCGGCGGTGTCAGAGTGGCGGTGAGACCG  
150 160 170 180 190 200  
220 X  
ACGACAT  
|||||  
ACGACAT  
X

ACGACAT  
|||||  
ACGACAT  
X

## 16. US-09-697-123B-9 (1-223)

US-09-697-123B-2 Sequence 24, Application US/09697123B

Initial Score = 136 Optimized Score = 160 Significance = 0.20  
Residue Identity = 76% Matches = 170 Mismatches = 38  
Gaps = 15 Conservative Substitutions = 0

X  
10 20 30 40 50 60 70  
TCAAGAGAAAGCCCTACGACCTGTCACGCTGACGAGCAAGGTCACAAAGTCAAGTGGGTCTGACGCGG  
|||||  
TCAAGAGAAAGCCCTACGACCTGTCACGCTGACGAGCAAGGTCACAAAGTCAAGTGGGTCTGACGCGG  
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140  
GCGAGCCGATCAGCTGTCACGCTGACGAGCAAGGTCACGAGCAAGTCAAGTGGGTCTGACGCGG  
|||||  
AGAAATGCGGCAACACACGACCTGACGAGCAAGGTCACGAGCAAGTCAAGTGGGTCTGACGCGG  
80 90 100 110 120 130 140

150 160 170 180 190 200 210  
ACCAAGCCGCTACGATGAGCCAGAGCCGCTGATGATGTCGCGGCAATCGAGTGGGTCTGACGCGG  
|||||  
ACGAGC-----GGCAAGCCGATGAGTCCCGGTCGAGTGGGTCTGACGCGG  
150 160 170 180 190 200

220 X  
ACGACAT  
|||||  
ACGACAT  
X

## 17. US-09-697-123B-9 (1-223)

US-09-697-123B-3 Sequence 3, Application US/09697123B

Initial Score = 133 Optimized Score = 171 Significance = 0.14  
Residue Identity = 82% Matches = 165 Mismatches = 23  
Gaps = 15 Conservative Substitutions = 0

X  
10 20 30 40 50 60 70  
TCAAGAGAAAGCCCTACGACCTGTCACGCTGACGAGCAAGGTCACAAAGTCAAGTGGGTCTGACGCGG  
|||||  
TCAAGAGAAAGCCCTACGACCTGTCACGCTGACGAGCAAGGTCACAAAGTCAAGTGGGTCTGACGCGG  
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140  
GCGAGCCGATCAGCTGTCACGCTGACGAGCAAGGTCACAAAGTCAAGTGGGTCTGACGCGG  
|||||  
GCGAGCCGATCAGCTGTCACGCTGACGAGCAAGGTCACAAAGTCAAGTGGGTCTGACGCGG  
80 90 100 110 120 130 140

150 160 170 180 190 200 210  
ACCAAGCCGCTACGATGAGCCAGAGCCGCTGATGATGTCGCGGCAATCGAGTGGGTCTGACGCGG  
|||||  
AC-----GAGGTCACGACAC---GATGACGTTCCGCGGCAACGAGTGGGTCTGACGCGG  
150 160 170 180 190 200

220 X  
ACGACAT  
|||||  
ACGACAT  
X

## 18. US-09-697-123B-9 (1-223)

US-09-697-123B-2 Sequence 20, Application US/09697123B

Initial Score = 131 Optimized Score = 163 Significance = 0.09  
Residue Identity = 82% Matches = 183 Mismatches = 22  
Gaps = 18 Conservative Substitutions = 0

X  
10 20 30 40 50 60 70  
TCAAGAGAAAGCCCTACGACCTGTCACGCTGACGAGCAAGGTCACAAAGTCAAGTGGGTCTGACGCGG  
|||||  
TCAAGAGAAAGCCCTACGACCTGTCACGCTGACGAGCAAGGTCACAAAGTCAAGTGGGTCTGACGCGG  
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140  
GCGAGCCGATCAGCTGTCACGCTGACGAGCAAGGTCACAAAGTCAAGTGGGTCTGACGCGG  
|||||  
GCGAGCCGATCAGCTGTCACGCTGACGAGCAAGGTCACAAAGTCAAGTGGGTCTGACGCGG  
80 90 100 110 120 130 140

150 160 170 180 190 200 210  
ACCAAGCCGCTACGATGAGCCAGAGCCGCTGATGATGTCGCGGCAATCGAGTGGGTCTGACGCGG  
|||||  
AC-----GAGGTCACGACAC---GATGACGTTCCCGGTCGAGTGGGTCTGACGCGG  
150 160 170 180 190 200

220 X  
ACGACAT  
|||||  
ACGACAT  
X

## 19. US-09-697-123B-9 (1-223)

US-09-697-123B-4 Sequence 4, Application US/09697123B

Initial Score = 127 Optimized Score = 163 Significance = 0.00  
Residue Identity = 79% Matches = 178 Mismatches = 29  
Gaps = 16 Conservative Substitutions = 0

X  
10 20 30 40 50 60 70  
TCAAGAGAAAGCCCTACGACCTGTCACGCTGACGAGCAAGGTCACAAAGTCAAGTGGGTCTGACGCGG  
|||||  
TCAAGAGAAAGCCCTACGACCTGTCACGCTGACGAGCAAGGTCACAAAGTCAAGTGGGTCTGACGCGG  
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140  
GCGAGCCGATCAGCTGTCACGCTGACGAGCAAGGTCACAAAGTCAAGTGGGTCTGACGCGG  
|||||  
GCGAGCCGATCAGCTGTCACGCTGACGAGCAAGGTCACAAAGTCAAGTGGGTCTGACGCGG  
80 90 100 110 120 130 140

150 160 170 180 190 200 210  
ACCAAGCCGCTACGATGAGCCAGAGCCGCTGATGATGTCGCGGCAATCGAGTGGGTCTGACGCGG  
|||||  
AC-----GAGGTCACGACACGATGA---GATGACGTTCCCGGTCGAGTGGGTCTGACGCGG  
150 160 170 180 190 200

220 X  
ACGACAT  
|||||  
ACGACAT  
X

## 20. US-09-697-123B-9 (1-223)

US-09-697-123B-1 Sequence 18, Application US/09697123B

Initial Score = 116 Optimized Score = 150 Significance = -0.25  
Residue Identity = 73% Matches = 165 Mismatches = 43  
Gaps = 18 Conservative Substitutions = 0

X  
10 20 30 40 50 60 70  
TCAAGAGAAAGCCCTACGACCTGTCACGCTGACGAGCAAGGTCACAAAGTCAAGTGGGTCTGACGCGG  
|||||  
TCAAGAGAAAGCCCTACGACCTGTCACGCTGACGAGCAAGGTCACAAAGTCAAGTGGGTCTGACGCGG  
X 10 20 30 40 50 60 70

140 150 160 170 180 190 200 210  
GCCTGACCAACGCCCGTACGATGGCCACCCCGCTCATCTGTCCTCCCGGGCATTCGAGGTGCCCGTGG

220 X  
AGACCGACGACAT  
|||||  
AGACCGACGACAT  
210 X

25. US-09-697-123B-9 (1-223)  
US-09-697-123B-2 Sequence 25, Application US/09697123B  
Initial Score = 19 Optimized Score = 19 Significance = -2.45  
Residue Identity = 100% Matches = 19 Mismatches = 0  
Gaps = 0 Conservative Substitutions = 0  
X 10 X  
TCACGACGAAAGCCCTACGA  
|||||  
TCACGACGAAAGCCCTACGA  
X 10 X



FastDB - Fast Pairwise Comparison of Sequences  
Release 5.4

Query sequence being compared:	US-09-697-123B-8 (1-208)
Number of sequences searched:	26
Number of scores above cutoff:	25

Results of the initial comparison of US-09-697-123B-8 (1-208) with:  
File : US09697123B.seq

[illegible]

PARAMETERS		
Similarity matrix	Unitary	4
Mismatch penalty	1	30
Gap penalty	5.00	30
Gap size penalty	0.33	Window size
Cutoff score	12	208
Randomization group	0	

## SEARCH STATISTICS

Scores:	Mean	Median	Standard Deviation
---------	------	--------	--------------------

	CPU	Total Elapsed
Times:		

00:00:00:00

Number of residues:	5077
Number of sequences searched:	26
Number of scores above cutoff:	25

The scores below are sorted by initial score. Significance is calculated based on initial score.

A 100% identical sequence to the query sequence was found:

Sequence Name	Description	Length	Int. Opt. Score	Sig. Frame
---------------	-------------	--------	-----------------	------------

Sequence Name	Description	Length	Score	Sig.	Frame
---------------	-------------	--------	-------	------	-------

2.	US-09-697-123B-1	Sequence 17, Application	208	187	187	0	68	0
3.	US-09-697-123B-1	Sequence 15, Application	208	185	185	0	64	0
4.	US-09-697-123B-1	Sequence 11, Application	208	184	185	0	62	0
5.	US-09-697-123B-3	Sequence 3, Application U	208	184	184	0	62	0
6.	US-09-697-123B-1	Sequence 19, Application	208	183	183	0	60	0
7.	US-09-697-123B-1	Sequence 1, Application U	208	183	183	0	60	0
8.	US-09-697-123B-1	Sequence 16, Application	208	182	182	0	58	0
9.	US-09-697-123B-1	Sequence 10, Application	208	182	182	0	58	0
10.	US-09-697-123B-6	Sequence 6, Application U	208	182	182	0	58	0
11.	US-09-697-123B-5	Sequence 5, Application U	208	182	182	0	58	0
12.	US-09-697-123B-2	Sequence 22, Application	208	181	181	0	55	0
13.	US-09-697-123B-2	Sequence 24, Application	208	179	179	0	51	0
14.	US-09-697-123B-2	Sequence 2, Application	208	178	178	0	49	0
15.	US-09-697-123B-2	Sequence 23, Application	208	177	177	0	47	0
16.	US-09-697-123B-2	Sequence 20, Application	205	163	173	0	18	0
17.	US-09-697-123B-1	Sequence 13, Application	203	154	161	0	00	0
18.	US-09-697-123B-1	Sequence 12, Application	207	153	174	-0.02	0	0
19.	US-09-697-123B-4	Sequence 4, Application	207	152	179	-0.04	0	0
20.	US-09-697-123B-9	Sequence 9, Application U	223	158	155	-0.33	0	0
21.	US-09-697-123B-7	Sequence 7, Application U	214	127	164	-0.55	0	0
22.	US-09-697-123B-2	Sequence 21, Application	214	122	170	-0.66	0	0
23.	US-09-697-123B-1	Sequence 14, Application	214	118	168	-0.74	0	0
24.	US-09-697-123B-1	Sequence 18, Application	211	117	164	-0.76	0	0
**** 2 standard deviations below mean ****			19	19	19	-2.77	0	0
25.	US-09-697-123B-2	Sequence 25, Application	19	19	19	-2.77	0	0

1. US-09-697-123B-8 (1-208)  
US-09-697-123B-8 Sequence 8, Application US/09697123B

Initial Score	-	208	Optimized Score	-	208	Significance	-	1.11
Residue Identity	-	100%	Matches	-	208	Mismatches	-	0
Gaps	-	0	Conservative Substitutions	-			-	0

X 10 20 30 40 50 60 70  
TCAAGGAGAACGCTACGACCTGGCCGCTGTCGGCGATACAGTCAACGAAAGCTGGCCCTGAACCC  
TCAAGGAGAACGCTACGACCTGGCCGCTGTCGGCGATACAGTCAACGAAAGCTGGCCCTGAACCC  
TCAAGGAGAACGCTACGACCTGGCCGCTGTCGGCGATACAGTCAACGAAAGCTGGCCCTGAACCC

ATCATCGATCCACGACGACGCTGACCGAAGACGTCGCGCACCATCGAGTATCTGTCCGCTGG  
80 90 100 110 120 130 140  
ATCATCGATCCACGACGACGCTGACCGAAGACGTCGCGCACCATCGAGTATCTGTCCGCTGG  
80 90 100 110 120 130 140

ACGAGGGCCAGGGCCACGAGTACCGTCCGGGGGGGTGAGTCCGGTGAACCCGACGACATAT  
150 160 170 180 190 200 X  
ACGAGGGCCAGGGCCACGAGTACCGTCCGGGGGGGTGAGTCCGGTGAACCCGACGACATAT  
150 160 170 180 190 200 X

2. US-09-697-123B-8 (1-208)  
US-09-697-123B-1 Sequence 17, Application US/09697123B

Initial Score	-	187	Optimized Score	-	187	Significance	-	0.68
Residue Identity	-	898	Matches	-	187	Mismatches	-	211
Gaps	-	0	Conservative Substitutions				-	0

X	10	20	30	40	50	60	70
TCAA	GGAAGCCCTAG	ACTGGCCGTG	TCGCGCGG	ATACAAGT	CAACAAGAA	GTGGGCTT	CAACACA
TCAA	GGAAGCCCTAG	ACTGGCCGTG	TCGCGCGG	ATACAAGT	CAACAAGAA	GTGGGCTT	CAACACA
TCAA	GGAAGCCCTAG	ACTGGCCGTG	TCGCGCGG	ATACAAGT	CAACAAGAA	GTGGGCTT	CAACACA
TCAA	GGAAGCCCTAG	ACTGGCCGTG	TCGCGCGG	ATACAAGT	CAACAAGAA	GTGGGCTT	CAACACA
X	10	20	30	40	50	60	70

x	10	20	30	40	50	60	70
---	----	----	----	----	----	----	----

8. US-09-697-123B-8 (1-208)  
US-09-697-123B-1 Sequence 16, Application US/09697123B

Initial Score = 182 Optimized Score = 182 Significance = 0.58  
Residue Identity = 87% Matches = 182 Mismatches = 26  
Gaps = 0 Conservative Substitutions = 0

X  
10 20 30 40 50 60 70  
TCAAGAGAACCGCTACGACCTGGCCCGCTGCGCCGATCAAGATCAAGAAAGAGCTGCGCTGAACACCA  
TCAAGAGAACCGCTACGACCTGGCCCGCTGCGCTGCTATAGGTCAACAGAAAGCTGCGCTGCTGCTGCG  
X 10 20 30 40 50 60 70

ATATCCCATACACCAACGACGCTGCGCCGCTGCGCCGATCAAGATCAAGAAAGAGCTGCGCTGCGCTGCG  
GCGACCCCATACGCTGCGACGCTGCGCCGCTGCGCCGATCAAGATCAAGAAAGAGCTGCGCTGCGCTGCG  
X 80 90 100 110 120 130 140

ACGAGGCCAGGACGACGCTGCGCCGCTGCGCCGATCAAGATCAAGAAAGAGCTGCGCTGCGCTGCGCTGCG  
ACGAGGCCAGGACGACGCTGCGCCGCTGCGCCGATCAAGATCAAGAAAGAGCTGCGCTGCGCTGCGCTGCG  
X 150 160 170 180 190 200  
ACGAGGCCAGGACGACGCTGCGCCGCTGCGCCGATCAAGATCAAGAAAGAGCTGCGCTGCGCTGCGCTGCG

9. US-09-697-123B-8 (1-208)  
US-09-697-123B-1 Sequence 10, Application US/09697123B

Initial Score = 182 Optimized Score = 182 Significance = 0.58  
Residue Identity = 87% Matches = 182 Mismatches = 26  
Gaps = 0 Conservative Substitutions = 0

X  
10 20 30 40 50 60 70  
TCAAGAGAACCGCTACGACCTGGCCCGCTGCGCCGATCAAGATCAAGAAAGAGCTGCGCTGAACACCA  
TCAAGAGAACCGCTACGACCTGGCCCGCTGCGCCGATCAAGATCAAGAAAGAGCTGCGCTGAACACCGG  
X 10 20 30 40 50 60 70

ATCATCCCATACACCAACGACGCTGCGCCGCTGCGCCGATCAAGATCAAGAAAGAGCTGCGCTGCGCTGCG  
GCGACCCCATACGCTGCGACGCTGCGCCGCTGCGCCGATCAAGATCAAGAAAGAGCTGCGCTGCGCTGCG  
X 80 90 100 110 120 130 140

ACGAGGCCAGGACGACGCTGCGCCGCTGCGCCGATCAAGATCAAGAAAGAGCTGCGCTGCGCTGCGCTGCG  
ACGAGGCCAGGACGACGCTGCGCCGCTGCGCCGATCAAGATCAAGAAAGAGCTGCGCTGCGCTGCGCTGCG  
X 150 160 170 180 190 200  
ACGAGGCCAGGACGACGCTGCGCCGCTGCGCCGATCAAGATCAAGAAAGAGCTGCGCTGCGCTGCGCTGCG

10. US-09-697-123B-8 (1-208)  
US-09-697-123B-6 Sequence 6, Application US/09697123B

Initial Score = 182 Optimized Score = 182 Significance = 0.58  
Residue Identity = 87% Matches = 182 Mismatches = 26  
Gaps = 0 Conservative Substitutions = 0

X  
10 20 30 40 50 60 70  
TCAAGAGAACCGCTACGACCTGGCCCGCTGCGCCGATCAAGATCAAGAAAGAGCTGCGCTGAACACCA  
TCAAGAGAACCGCTACGACCTGGCCCGCTGCGCCGATCAAGATCAAGAAAGAGCTGCGCTGAACACCGG  
X 10 20 30 40 50 60 70

ATCATCCCATACACCAACGACGCTGCGCCGCTGCGCCGATCAAGATCAAGAAAGAGCTGCGCTGCGCTGCG  
GCGACCCCATACGCTGCGACGCTGCGCCGCTGCGCCGATCAAGATCAAGAAAGAGCTGCGCTGCGCTGCG  
X 80 90 100 110 120 130 140

ACGAGGCCAGGACGACGCTGCGCCGCTGCGCCGATCAAGATCAAGAAAGAGCTGCGCTGCGCTGCGCTGCG  
ACGAGGCCAGGACGACGCTGCGCCGCTGCGCCGATCAAGATCAAGAAAGAGCTGCGCTGCGCTGCGCTGCG  
X 150 160 170 180 190 200  
ACGAGGCCAGGACGACGCTGCGCCGCTGCGCCGATCAAGATCAAGAAAGAGCTGCGCTGCGCTGCGCTGCG

11. US-09-697-123B-8 (1-208)  
US-09-697-123B-5 Sequence 5, Application US/09697123B

Initial Score = 182 Optimized Score = 182 Significance = 0.58  
Residue Identity = 87% Matches = 182 Mismatches = 26  
Gaps = 0 Conservative Substitutions = 0

X  
10 20 30 40 50 60 70  
TCAAGAGAACCGCTACGACCTGGCCCGCTGCGCCGATCAAGATCAAGAAAGAGCTGCGCTGAACACCA  
TCAAGAGAACCGCTACGACCTGGCCCGCTGCGCCGATCAAGATCAAGAAAGAGCTGCGCTGAACACCGG  
X 10 20 30 40 50 60 70

ATATCCCATACACCAACGACGCTGCGCCGCTGCGCCGATCAAGATCAAGAAAGAGCTGCGCTGCGCTGCG  
GCGACCCCATACGCTGCGACGCTGCGCCGCTGCGCCGATCAAGATCAAGAAAGAGCTGCGCTGCGCTGCG  
X 80 90 100 110 120 130 140

ACGAGGCCAGGACGACGCTGCGCCGCTGCGCCGATCAAGATCAAGAAAGAGCTGCGCTGCGCTGCGCTGCG  
ACGAGGCCAGGACGACGCTGCGCCGCTGCGCCGATCAAGATCAAGAAAGAGCTGCGCTGCGCTGCGCTGCG  
X 150 160 170 180 190 200  
ACGAGGCCAGGACGACGCTGCGCCGCTGCGCCGATCAAGATCAAGAAAGAGCTGCGCTGCGCTGCGCTGCG

12. US-09-697-123B-8 (1-208)  
US-09-697-123B-2 Sequence 22, Application US/09697123B

Initial Score = 181 Optimized Score = 181 Significance = 0.55  
Residue Identity = 87% Matches = 181 Mismatches = 27  
Gaps = 0 Conservative Substitutions = 0

X  
10 20 30 40 50 60 70  
TCAAGAGAACCGCTACGACCTGGCCCGCTGCGCCGATCAAGATCAAGAAAGAGCTGCGCTGAACACCA  
TCAAGAGAACCGCTACGACCTGGCCCGCTGCGCCGATCAAGATCAAGAAAGAGCTGCGCTGAACACCGG  
X 10 20 30 40 50 60 70

ATCATCCCATACACCAACGACGCTGCGCCGCTGCGCCGATCAAGATCAAGAAAGAGCTGCGCTGCGCTGCG  
GCGACCCCATACGCTGCGACGCTGCGCCGCTGCGCCGATCAAGATCAAGAAAGAGCTGCGCTGCGCTGCG  
X 80 90 100 110 120 130 140

ACGAGGCCAGGACGACGCTGCGCCGCTGCGCCGATCAAGATCAAGAAAGAGCTGCGCTGCGCTGCGCTGCG  
ACGAGGCCAGGACGACGCTGCGCCGCTGCGCCGATCAAGATCAAGAAAGAGCTGCGCTGCGCTGCGCTGCG  
X 150 160 170 180 190 200  
ACGAGGCCAGGACGACGCTGCGCCGCTGCGCCGATCAAGATCAAGAAAGAGCTGCGCTGCGCTGCGCTGCG

13. US-09-697-123B-8 (1-208)  
US-09-697-123B-2 Sequence 24, Application US/09697123B

Initial Score = 179 Optimized Score = 179 Significance = 0.51  
Residue Identity = 86% Matches = 179 Mismatches = 29  
Gaps = 0 Conservative Substitutions = 0

X  
10 20 30 40 50 60 70  
TCAAGAGAACCGCTACGACCTGGCCCGCTGCGCCGATCAAGATCAAGAAAGAGCTGCGCTGAACACCA  
TCAAGAGAACCGCTACGACCTGGCCCGCTGCGCCGATCAAGATCAAGAAAGAGCTGCGCTGAACACCGG  
X 10 20 30 40 50 60 70

ATCATCCCATACACCAACGACGCTGCGCCGCTGCGCCGATCAAGATCAAGAAAGAGCTGCGCTGCGCTGCG  
AGATGCGCCCAACACGACGCTGCGCCGCTGCGCCGATCAAGATCAAGAAAGAGCTGCGCTGCGCTGCGCTGCG  
X 80 90 100 110 120 130 140

80 90 100 110 120 130 140

US-09-697-123B-4 Sequence 4, Application US/09697123B

Initial Score = 152 Optimized Score = 179 Significance = -0.04  
Residue Identity = 88% Matches = 184 Mismatches = 23  
Gaps = 1 Conservative Substitutions = 0

X 10 20 30 40 50 60 70  
TCACGAGAAAGGCGTACGACCTGGCCCGCTGCGCCGATACAGGTCAACAAGAAAGCTGGCCGTAACACCA  
TCACGAGAAAGGCGTACGACCTGGCCCGCTGCGCCGATACAGGTCAACAAGAAAGCTGGCCGCTGATGTCG  
X 10 20 30 40 50 60 70

ATCATCTCATACACGACGACGCTGCTGCGCCGATACAGGTCAACAAGAAAGCTGGCCGCTGATGTCG  
CGATCCCATACACGACGCTGCTGCGCCGATACAGGTCAACAAGAAAGCTGGCCGCTGATGTCG  
X 10 20 30 40 50 60 70

150 160 170 180 190 200  
ACGAGCGCCAGCGACGATGACCTGGCCCGCTGCGCCGATACAGGTCAACAAGAAAGCTGGCCGCTGATGTCG  
ACGAGCGCGACGACGATGACCTGGCCCGCTGCGCCGATACAGGTCAACAAGAAAGCTGGCCGCTGATGTCG  
X 150 160 170 180 190 200

20. US-09-697-123b-8 (1-208)  
US-09-697-123b-9 Sequence 9, Application US/09697123b

Initial Score = 138 Optimized Score = 165 Significance = -0.33  
Residue Identity = 78% Matches = 175 Mismatches = 33  
Gaps = 15 Conservative Substitutions = 0

X 10 20 30 40 50 60 70  
TCACGAGAAAGGCGTACGACCTGGCCCGCTGCGCCGATACAGGTCAACAAGAAAGCTGGCCGCTGATGTCG  
TCACGAGAAAGGCGTACGACCTGGCCCGCTGCGCCGATACAGGTCAACAAGAAAGCTGGCCGCTGATGTCG  
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140  
ATCATCTCATACACGACGACGCTGCTGCGCCGATACAGGTCAACAAGAAAGCTGGCCGCTGATGTCG  
CGAGTCGCGCTACAGGTACGACCTGGCCCGCTGCGCCGATACAGGTCAACAAGAAAGCTGGCCGCTGATGTCG  
X 80 90 100 110 120 130 140

150 160 170 180 190 200  
ACGAGCGCCAGCGACGATGACCTGGCCCGCTGCGCCGATACAGGTCAACAAGAAAGCTGGCCGCTGATGTCG  
ACGAGCGCGACGACGATGACCTGGCCCGCTGCGCCGATACAGGTCAACAAGAAAGCTGGCCGCTGATGTCG  
X 150 160 170 180 190 200

ACGAGAT 111111  
ACGAGAT 111111  
X 220 X

21. US-09-697-123b-8 (1-208)  
US-09-697-123b-7 Sequence 7, Application US/09697123b

Initial Score = 127 Optimized Score = 164 Significance = -0.55  
Residue Identity = 79% Matches = 171 Mismatches = 37  
Gaps = 6 Conservative Substitutions = 0

X 10 20 30 40 50 60 70  
TCACGAGAAAGGCGTACGACCTGGCCCGCTGCGCCGATACAGGTCAACAAGAAAGCTGGCCGCTGATGTCG  
TCACGAGAAAGGCGTACGACCTGGCCCGCTGCGCCGATACAGGTCAACAAGAAAGCTGGCCGCTGATGTCG  
X 10 20 30 40 50 60 70

70 80 90 100 110 120 130  
ACACCAATCATCATACGACGACGCTGCTGCGCCGATACAGGTCAACAAGAAAGCTGGCCGCTGATGTCG  
CCACCGCGCTGCTGATGTCGACGACGCTGCTGCGCCGATACAGGTCAACAAGAAAGCTGGCCGCTGATGTCG  
X 80 90 100 110 120 130 140

140 150 160 170 180 190 200  
GCCCTCAGAGAGGCGCCAGCGACGATGACCTGGCCCGCTGCGCCGATACAGGTCAACAAGAAAGCTGGCCGCTGATGTCG  
GCCCTCAGAGAGGCGCCAGCGACGATGACCTGGCCCGCTGCGCCGATACAGGTCAACAAGAAAGCTGGCCGCTGATGTCG  
X 140 150 160 170 180 190 200

22. US-09-697-123b-8 (1-208)  
US-09-697-123b-2 Sequence 21, Application US/09697123b

Initial Score = 122 Optimized Score = 170 Significance = -0.66  
Residue Identity = 82% Matches = 177 Mismatches = 31  
Gaps = 6 Conservative Substitutions = 0

X 10 20 30 40 50 60 70  
TCACGAGAAAGGCGTACGACCTGGCCCGCTGCGCCGATACAGGTCAACAAGAAAGCTGGCCGCTGATGTCG  
TCACGAGAAAGGCGTACGACCTGGCCCGCTGCGCCGATACAGGTCAACAAGAAAGCTGGCCGCTGATGTCG  
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140  
ATCATCTCATACACGACGACGCTGCTGCGCCGATACAGGTCAACAAGAAAGCTGGCCGCTGATGTCG  
CCATTCGCGCTACAGGTACGACCTGGCCCGCTGCGCCGATACAGGTCAACAAGAAAGCTGGCCGCTGATGTCG  
X 80 90 100 110 120 130 140

140 150 160 170 180 190 200  
GCCCTCAGAGAGGCGCCAGCGACGATGACCTGGCCCGCTGCGCCGATACAGGTCAACAAGAAAGCTGGCCGCTGATGTCG  
GCCCTCAGAGAGGCGCCAGCGACGATGACCTGGCCCGCTGCGCCGATACAGGTCAACAAGAAAGCTGGCCGCTGATGTCG  
X 140 150 160 170 180 190 200

23. US-09-697-123b-8 (1-208)  
US-09-697-123b-1 Sequence 14, Application US/09697123b

Initial Score = 118 Optimized Score = 168 Significance = -0.74  
Residue Identity = 81% Matches = 175 Mismatches = 33  
Gaps = 6 Conservative Substitutions = 0

X 10 20 30 40 50 60 70  
TCACGAGAAAGGCGTACGACCTGGCCCGCTGCGCCGATACAGGTCAACAAGAAAGCTGGCCGCTGATGTCG  
TCACGAGAAAGGCGTACGACCTGGCCCGCTGCGCCGATACAGGTCAACAAGAAAGCTGGCCGCTGATGTCG  
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140  
ATCATCTCATACACGACGACGCTGCTGCGCCGATACAGGTCAACAAGAAAGCTGGCCGCTGATGTCG  
CCAGTCGCGCTACAGGTACGACCTGGCCCGCTGCGCCGATACAGGTCAACAAGAAAGCTGGCCGCTGATGTCG  
X 80 90 100 110 120 130 140

140 150 160 170 180 190 200  
GCCCTCAGAGAGGCGCCAGCGACGATGACCTGGCCCGCTGCGCCGATACAGGTCAACAAGAAAGCTGGCCGCTGATGTCG  
GCCCTCAGAGAGGCGCCAGCGACGATGACCTGGCCCGCTGCGCCGATACAGGTCAACAAGAAAGCTGGCCGCTGATGTCG  
X 140 150 160 170 180 190 200

24. US-09-697-123b-8 (1-208)  
US-09-697-123b-1 Sequence 18, Application US/09697123b

Initial Score = 117 Optimized Score = 164 Significance = -0.76  
Residue Identity = 80% Matches = 170 Mismatches = 38  
Gaps = 3 Conservative Substitutions = 0

X 10 20 30 40 50 60 70  
TCACGAGAAAGGCGTACGACCTGGCCCGCTGCGCCGATACAGGTCAACAAGAAAGCTGGCCGCTGATGTCG  
TCACGAGAAAGGCGTACGACCTGGCCCGCTGCGCCGATACAGGTCAACAAGAAAGCTGGCCGCTGATGTCG  
X 10 20 30 40 50 60 70



FastDB - Fast Pairwise Comparison of Sequences  
Release 5.4

Query sequence being compared:	US-09-697-123B-7 (1-214)
Number of sequences searched:	26
Number of scores above cutoff:	25

Results of the initial comparison of US-09-697-123B-7 (1-214) with:  
File : US09697123B.seq

The following table summarizes the data points shown in the dot plot:

Category	Score	Stdev
N	100	0
U	50	0
M	-2	0
B	-1	0
E	-1	0
R	-1	0
O	-1	0
F	-1	0
S	-1	0
E	-1	0
O	-1	0
U	-1	0
E	-1	0
N	-1	0
C	-1	0
E	-1	0
S	-1	0
S	-1	0
E	-1	0
O	-1	0
U	-1	0
E	-1	0
N	-1	0
C	-1	0
E	-1	0
S	-1	0
S	-1	0
E	-1	0
O	-1	0
U	-1	0
E	-1	0
N	-1	0
C	-1	0
E	-1	0
S	-1	0
S	-1	0
E	-1	0
O	-1	0
U	-1	0
E	-1	0
N	-1	0
C	-1	0
E	-1	0
S	-1	0
S	-1	0
E	-1	0
O	-1	0
U	-1	0
E	-1	0
N	-1	0
C	-1	0
E	-1	0
S	-1	0
S	-1	0
E	-1	0
O	-1	0
U	-1	0
E	-1	0
N	-1	0
C	-1	0
E	-1	0
S	-1	0
S	-1	0
E	-1	0
O	-1	0
U	-1	0
E	-1	0
N	-1	0
C	-1	0
E	-1	0
S	-1	0
S	-1	0
E	-1	0
O	-1	0
U	-1	0
E	-1	0
N	-1	0
C	-1	0
E	-1	0
S	-1	0
S	-1	0
E	-1	0
O	-1	0
U	-1	0
E	-1	0
N	-1	0
C	-1	0
E	-1	0
S	-1	0
S	-1	0
E	-1	0
O	-1	0
U	-1	0
E	-1	0
N	-1	0
C	-1	0
E	-1	0
S	-1	0
S	-1	0
E	-1	0
O	-1	0
U	-1	0
E	-1	0
N	-1	0
C	-1	0
E	-1	0
S	-1	0
S	-1	0
E	-1	0
O	-1	0
U	-1	0
E	-1	0
N	-1	0
C	-1	0
E	-1	0
S	-1	0
S	-1	0
E	-1	0
O	-1	0
U	-1	0
E	-1	0
N	-1	0
C	-1	0
E	-1	0
S	-1	0
S	-1	0
E	-1	0
O	-1	0
U	-1	0
E	-1	0
N	-1	0
C	-1	0
E	-1	0
S	-1	0
S	-1	0
E	-1	0
O	-1	0
U	-1	0
E	-1	0
N	-1	0
C	-1	0
E	-1	0
S	-1	0
S	-1	0
E	-1	0
O	-1	0
U	-1	0
E	-1	0
N	-1	0
C	-1	0
E	-1	0
S	-1	0
S	-1	0
E	-1	0
O	-1	0
U	-1	0
E	-1	0
N	-1	0
C	-1	0
E	-1	0
S	-1	0
S	-1	0
E	-1	0
O	-1	0
U	-1	0
E	-1	0
N	-1	0
C	-1	0
E	-1	0
S	-1	0
S	-1	0
E	-1	0
O	-1	0
U	-1	0
E	-1	0
N	-1	0
C	-1	0
E	-1	0
S	-1	0

PARAMETERS		
Unary	K-tuple	4
Similarity matrix	Joining penalty	30
Mismatch penalty	Window size	208
Gap penalty		
Gap size penalty		
Cutoff score		
Randomization group		

Scores:	Mean	Median	Standard Deviation
	117	122	42.12
Times:	CPU		Total Elapsed
	00:00:00.00		00:00:00.00
Number of residues:		5077	
Number of sequences searched:		26	
Number of scores above cutoff:		25	

A 100% identical sequence to the query sequence was found.

Sequence Name	Description	Length	Init. Opt. Score	Sig. Frame
---------------	-------------	--------	------------------	------------

Sequence Name	Description	Intf. Opt.	Length	Score	Score	Sig.	Frame
---------------	-------------	------------	--------	-------	-------	------	-------

2.	US-09-697-123B-2	Sequence 21,	Application	214	197	1.90	0
3.	US-09-697-123B-1	Sequence 14,	Application	214	172	1.31	0
		**** 0 standard deviation from mean ****					
4.	US-09-697-123B-1	Sequence 17,	Application	208	137	0.47	0
5.	US-09-697-123B-1	Sequence 15,	Application	208	131	0.33	0
6.	US-09-697-123B-2	Sequence 2,	Application	208	129	0.28	0
7.	US-09-697-123B-1	Sequence 1,	Application	208	129	0.28	0
8.	US-09-697-123B-8	Sequence 8,	Application	208	127	0.24	0
9.	US-09-697-123B-3	Sequence 3,	Application	208	125	0.19	0
10.	US-09-697-123B-2	Sequence 24,	Application	208	123	0.14	0
11.	US-09-697-123B-1	Sequence 18,	Application	211	122	0.12	0
12.	US-09-697-123B-1	Sequence 16,	Application	208	121	0.09	0
13.	US-09-697-123B-6	Sequence 6,	Application	208	121	0.09	0
14.	US-09-697-123B-5	Sequence 5,	Application	208	121	0.09	0
15.	US-09-697-123B-2	Sequence 23,	Application	208	120	0.07	0
16.	US-09-697-123B-2	Sequence 22,	Application	208	117	0.00	0
17.	US-09-697-123B-1	Sequence 19,	Application	208	115	0.05	0
18.	US-09-697-123B-2	Sequence 20,	Application	205	109	0.15	0
19.	US-09-697-123B-1	Sequence 11,	Application	208	108	-0.21	0
20.	US-09-697-123B-1	Sequence 10,	Application	208	107	-0.24	0
21.	US-09-697-123B-1	Sequence 12,	Application	207	101	0.38	0
22.	US-09-697-123B-4	Sequence 4,	Application	207	92	-0.59	0
23.	US-09-697-123B-1	Sequence 13,	Application	223	92	-0.59	0
24.	US-09-697-123B-9	Sequence 9,	Application	223	91	-0.62	0
		**** 2 standard deviations below mean ****					
25.	US-09-697-123B-2	Sequence 25,	Application	19	19	-2.33	0

```

1. US-09-697-123B-7 (1-214)
   US-09-697-123B-7 Sequence 7, Application US/09697123B
Initial Score - 214 Optimized Score - 214 Significance - 2.30
Residue Identity - 100% Matches - 214 Mismatches - 0
Gaps - 0 Conservative Substitutions - 0

```

[illegible]

2. US-09-697-123B-7 (1-214)  
US-09-697-123B-2 Sequence 21, Application US/09697123B

	10	20	30	40	50	60	70
TCACAGAGAGGCGCTACGACCTGGCCCGCGGGCCGGTACAGGTCGACAGAGAACCTCGGCTTCGGCGGTCG							
TCACAGAGAGGCGCTACGACCTGGCCCGCGGGCCGGTACAGGTCGACAGAGAACCTCGGCTTCGGCGGCGCA							
Initial Score	-	197	Optimized score	-	197	Significance	- 1.90
Residue Identity	-	92%	Matches	-	197	Mismatches	- 17
Gaps	-	0	Conservative Substitutions	-			- 0

X 10 20 30 40 50 60 70

CCAACCCCGCTCTGTGTAAGTCCACACGCTCAACCGAGAGACGTGTCGCCACCATCGGTAACCTGGTGC  
|||||  
CCAAATCCGGCTGACGTGACCCACACCCCTCAACCGAGAAAGGTGTCGCCACCATGAGTACCTGGTGC  
80 90 100 110 120 130 140

GCCTCCACGAGGCGCCAGACCCACGATGACCCCGCCCGCCGCTTCGAGGTCCCGGTGAGGTGAGTGCACAT  
|||||  
GCCTCCACGAGGCGCCAGACCCACGATGACCCCGCCCGCCGCTTCGAGGTCCCGGTGAGGTGAGTGCACAT  
150 160 170 180 190 200 210 X

3. US-09-697-123B-7 (1-214)

US-09-697-123B-1 Sequence 14, Application US/09697123B

Initial Score = 172 Optimized Score = 172 Significance = 1.31  
Residue Identity = 80% Matches = 172 Mismatches = 42  
Gaps = 0 Conservative Substitutions = 0

X 10 20 30 40 50 60 70  
TCAAGAGAAACCGCTACGACCTGCGCCGCTGGCGCGGTCAAGGTGAACAGAAAGCTGGTCTTGCGGTG  
|||||  
TCAAGAGAAACCGCTACGACCTGCGCCGCTGGCGCGGTCAAGGTGAACAGAAAGCTGGTCTTGCGGTG  
10 20 30 40 50 60 70

CCAACCCCGCTCTGTGTAAGTCCACACGCTCAACCGAGAGACGTGTCGCCACCATCGGTAACCTGGTGC  
|||||  
CCGAGTCCGCGCTACCGCTGACCGCTGACCGAGCGGATGTCGTCGACCATCGGTAACCTGGTGC  
80 90 100 110 120 130 140

GCCTCCACGAGGCGCCAGACCCACGATGACCCCGCCCGCTTCGAGGTCCCGGTGAGGTGAGTGCACAT  
|||||  
GCCTCCACGAGGCGCCAGACCCACGATGACCCCGCCCGCTTCGAGGTCCCGGTGAGGTGAGTGCACAT  
150 160 170 180 190 200 210 X

4. US-09-697-123B-7 (1-214)

US-09-697-123B-1 Sequence 17, Application US/09697123B

Initial Score = 137 Optimized Score = 165 Significance = 0.47  
Residue Identity = 80% Matches = 172 Mismatches = 36  
Gaps = 6 Conservative Substitutions = 0

X 10 20 30 40 50 60 70  
TCAAGAGAAACCGCTACGACCTGCGCCGCTGGCGCGGTCAAGGTGAACAGAAAGCTGGTCTTGCGGTG  
|||||  
TCAAGAGAAACCGCTACGACCTGCGCCGCTGGCGCGGTCAAGGTGAACAGAAAGCTGGTCTTGCGGTG  
10 20 30 40 50 60 70

CCAACCCCGCTCTGTGTAAGTCCACACGCTCAACCGAGAGACGTGTCGCCACCATCGGTAACCTGGTGC  
|||||  
ACACCCGCTCTGTGTAAGTCCACACGCTCAACCGAGAGACGTGTCGCCACCATCGGTAACCTGGTGC  
70 80 90 100 110 120 130

GCCTCCACGAGGCGCCAGACCCACGATGACCCCGCCCGCTTCGAGGTCCCGGTGAGGTGAGTGCACAT  
|||||  
GCCTCCACGAGGCGCCAGACCCACGATGACCCCGCCCGCTTCGAGGTCCCGGTGAGGTGAGTGCACAT  
140 150 160 170 180 190 200 X

5. US-09-697-123B-7 (1-214)

US-09-697-123B-1 Sequence 15, Application US/09697123B

Initial Score = 131 Optimized Score = 170 Significance = 0.33  
Residue Identity = 82% Matches = 177 Mismatches = 31  
Gaps = 6 Conservative Substitutions = 0

X 10 20 30 40 50 60 70  
TCAAGAGAAACCGCTACGACCTGCGCCGCTGGCGCGGTCAAGGTGAACAGAAAGCTGGTCTTGCGGTG  
|||||  
TCAAGAGAAACCGCTACGACCTGCGCCGCTGGCGCGGTCAAGGTGAACAGAAAGCTGGTCTTGCGGTG  
X 10 20 30 40 50 60 70

CCAACCCCGCTCTGTGTAAGTCCACACGCTCAACCGAGAGACGTGTCGCCACCATCGGTAACCTGGTGC  
|||||  
GTGAGCC-----GATCACCAGCTCAACGCTGACCCAGAAAGACGTGTCGCCACCATCGGTAACCTGGTGC  
80 90 100 110 120 130 140

GCCTCCACGAGGCGCCAGACCCACGATGACCCCGCCCGCTTCGAGGTCCCGGTGAGGTGAGTGCACAT  
|||||  
GCCTCCACGAGGCGCCAGACCCACGATGACCCCGCCCGCTTCGAGGTCCCGGTGAGGTGAGTGCACAT  
140 150 160 170 180 190 200 X

6. US-09-697-123B-7 (1-214)

US-09-697-123B-2 Sequence 2, Application US/09697123B

Initial Score = 129 Optimized Score = 166 Significance = 0.28  
Residue Identity = 80% Matches = 173 Mismatches = 35  
Gaps = 6 Conservative Substitutions = 0

X 10 20 30 40 50 60 70  
TCAAGAGAAACCGCTACGACCTGCGCCGCTGGCGCGGTCAAGGTGAACAGAAAGCTGGTCTTGCGGTG  
|||||  
TCAAGAGAAACCGCTACGACCTGCGCCGCTGGCGCGGTCAAGGTGAACAGAAAGCTGGTCTTGCGGTG  
X 10 20 30 40 50 60 70

CCAACCCCGCTCTGTGTAAGTCCACACGCTCAACCGAGAGACGTGTCGCCACCATCGGTAACCTGGTGC  
|||||  
GCAAGCC-----GATCACCAGCTCAACGCTGACCCAGAGAAAGCTGTAACCATCGGTAACCTGGTGC  
80 90 100 110 120 130 140

GCCTCCACGAGGCGCCAGACCCACGATGACCCCGCCCGCTTCGAGGTCCCGGTGAGGTGAGTGCACAT  
|||||  
GCCTCCACGAGGCGCCAGACCCACGATGACCCCGCCCGCTTCGAGGTCCCGGTGAGGTGAGTGCACAT  
140 150 160 170 180 190 200 X

7. US-09-697-123B-7 (1-214)

US-09-697-123B-1 Sequence 1, Application US/09697123B

Initial Score = 129 Optimized Score = 168 Significance = 0.28  
Residue Identity = 81% Matches = 175 Mismatches = 33  
Gaps = 6 Conservative Substitutions = 0

X 10 20 30 40 50 60 70  
TCAAGAGAAACCGCTACGACCTGCGCCGCTGGCGCGGTCAAGGTGAACAGAAAGCTGGTCTTGCGGTG  
|||||  
TCAAGAGAAACCGCTACGACCTGCGCCGCTGGCGCGGTCAAGGTGAACAGAAAGCTGGTCTTGCGGTG  
X 10 20 30 40 50 60 70

CCAACCCCGCTCTGTGTAAGTCCACACGCTCAACCGAGAGACGTGTCGCCACCATCGGTAACCTGGTGC  
|||||  
GC-----GATCACCAGCTCAACGCTGACCCAGAGAGACGTGTCGCCACCATCGGTAACCTGGTGC  
80 90 100 110 120 130 140

GCCTCCACGAGGCGCCAGACCCACGATGACCCCGCCCGCTTCGAGGTCCCGGTGAGGTGAGTGCACAT  
|||||  
GCCTCCACGAGGCGCCAGACCCACGATGACCCCGCCCGCTTCGAGGTCCCGGTGAGGTGAGTGCACAT  
140 150 160 170 180 190 200 X

8. US-09-697-123B-7 (1-214)



## US-09-697-123B-8 Sequence 8, Application US/09697123B

Initial Score = 127 Optimized Score = 164 Significance = 0.24  
Residue Identity = 79% Matches = 171 Mismatches = 37  
Gaps = 6 Conservative Substitutions = 0

X 10 20 30 40 50 60 70  
TCAGAGGAAGGCGTACGACCGCGCCCGCTGGCGGTACAGAGTGAACAAAGAGTGGTCTTGGCGGTG  
|||||  
TCAGAGGAAGGCGTACGACCGCGCCCGCTGGCGGTACAGAGTGAACAAAGAGTGGTCTTGGCGGTG  
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140  
CCACCCGCGCTGTGGTGCACGACCGCTACCGACGAGAGAGAGTGGTCCGACATCGGATACCTGTC  
|||||  
ACACCAATTCATCCGATACACGAGAGAGCTGACCGAGAGAGAGTGGTCCGACATCGGATACCTGTC  
70 80 90 100 110 120 130

150 160 170 180 190 200 210 X  
GCTTCACAGAGGCGGACGACGATGACCGCGCGCGCGCTGAGAGTCCCGGTGAGAGTGCACGACAT  
|||||  
GCTTCACAGAGGCGGACGACGATGACCGCGCGCGCGCTGAGAGTCCCGGTGAGAGTGCACGACAT  
140 150 160 170 180 190 200 X

9. US-09-697-123B-7 (1-214)  
US-09-697-123B-3 Sequence 3, Application US/09697123B

Initial Score = 125 Optimized Score = 164 Significance = 0.19  
Residue Identity = 79% Matches = 171 Mismatches = 37  
Gaps = 6 Conservative Substitutions = 0

X 10 20 30 40 50 60 70  
TCAGAGGAAGGCGTACGACCGCGCGCGCTGAGAGTGAACAAAGAGTGGTCTTGGCGGTG  
|||||  
TCAGAGGAAGGCGTACGACCGCGCGCGCTGAGAGTGAACAAAGAGTGGTCTTGGCGGTG  
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140  
CCACCCGCGCTGTGGTGCACGACCGCTACCGACGAGAGAGTGGTCCGACATCGGATACCTGTC  
|||||  
GC-----GATCCGATACCGACGCTCCACGCTGACCGAGAGAGAGTGGTCCGACATCGGATACCTGTC  
80 90 100 110 120 130

150 160 170 180 190 200 210 X  
GCTTCACAGAGGCGGACGACGATGACCGCGCGCGCTGAGAGTGGTCTTGGCGGTG  
|||||  
GCTTCACAGAGGCGGACGACGATGACCGCGCGCGCTGAGAGTGGTCTTGGCGGTG  
140 150 160 170 180 190 200 X

10. US-09-697-123B-7 (1-214)  
US-09-697-123B-2 Sequence 24, Application US/09697123B

Initial Score = 123 Optimized Score = 161 Significance = 0.14  
Residue Identity = 78% Matches = 168 Mismatches = 40  
Gaps = 6 Conservative Substitutions = 0

X 10 20 30 40 50 60 70  
TCAGAGGAAGGCGTACGACCGCGCGCGCTGAGAGTGAACAAAGAGTGGTCTTGGCGGTG  
|||||  
TCAGAGGAAGGCGTACGACCGCGCGCGCTGAGAGTGAACAAAGAGTGGTCTTGGCGGTG  
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140  
CCACCCGCGCTGTGGTGCACGACCGCTACCGACGAGAGAGTGGTCCGACATCGGATACCTGTC  
|||||  
AGATAGGCGC-----AACACGACGACCGCTGACCGAGAGAGAGTGGTCCGACATCGGATACCTGTC  
80 90 100 110 120 130

150 160 170 180 190 200 210 X  
GCTTCACAGAGGCGGACGACGATGACCGCGCGCGCTGAGAGTCCCGGTGAGAGTGCACGACAT  
|||||  
GCTTCACAGAGGCGGACGACGATGACCGCGCGCGCTGAGAGTCCCGGTGAGAGTGCACGACAT  
150 160 170 180 190 200 210 X

GCTTCACAGAGGCGGACGACGATGAGAGTCCCGCGGTGAGAGTGGTGGTGAACGACGACAT  
140 150 160 170 180 190 200 X

11. US-09-697-123B-7 (1-214)  
US-09-697-123B-1 Sequence 18, Application US/09697123B

Initial Score = 122 Optimized Score = 159 Significance = 0.12  
Residue Identity = 81% Matches = 175 Mismatches = 36  
Gaps = 3 Conservative Substitutions = 0

X 10 20 30 40 50 60 70  
TCAGAGGAAGGCGTACGACCGCGCGCGCTGAGAGTGAACAAAGAGTGGTCTTGGCGGTG  
|||||  
TCAGAGGAAGGCGTACGACCGCGCGCGCTGAGAGTGAACAAAGAGTGGTCTTGGCGGTG  
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140  
CCACCCGCGCTGTGGTGCACGACCGCTACCGACGAGAGAGTGGTCCGACATCGGATACCTGTC  
|||||  
AGAACCCGCGCGACACGACGCTGACGACGAGAGAGAGTGGTCCGACATCGGATACCTGTC  
70 80 90 100 110 120 130 140

150 160 170 180 190 200 210 X  
GCTTCACAGAGGCGGACGACGATGACCGCGCGCGCTGAGAGTGGTCTTGGCGGTG  
|||||  
GCTTCACAGAGGCGGACGACGATGACCGCGCGCGCTGAGAGTGGTCTTGGCGGTG  
150 160 170 180 190 200 210 X

12. US-09-697-123B-7 (1-214)  
US-09-697-123B-1 Sequence 16, Application US/09697123B

Initial Score = 121 Optimized Score = 159 Significance = 0.09  
Residue Identity = 77% Matches = 166 Mismatches = 42  
Gaps = 6 Conservative Substitutions = 0

X 10 20 30 40 50 60 70  
TCAGAGGAAGGCGTACGACCGCGCGCGCTGAGAGTGAACAAAGAGTGGTCTTGGCGGTG  
|||||  
TCAGAGGAAGGCGTACGACCGCGCGCGCTGAGAGTGAACAAAGAGTGGTCTTGGCGGTG  
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140  
CCACCCGCGCTGTGGTGCACGACCGCTACCGACGAGAGAGTGGTCCGACATCGGATACCTGTC  
|||||  
GCGAGC-----CATCAGCTCGGTCAGCGCTGACGAGAGAGAGTGGTCCGACATCGGATACCTGTC  
80 90 100 110 120 130

150 160 170 180 190 200 210 X  
GCTTCACAGAGGCGGACGACGATGACCGCGCGCGCTGAGAGTGGTCTTGGCGGTG  
|||||  
GCTTCACAGAGGCGGACGACGATGACCGCGCGCGCTGAGAGTGGTCTTGGCGGTG  
140 150 160 170 180 190 200 X

13. US-09-697-123B-7 (1-214)  
US-09-697-123B-6 Sequence 6, Application US/09697123B

Initial Score = 121 Optimized Score = 159 Significance = 0.09  
Residue Identity = 77% Matches = 166 Mismatches = 42  
Gaps = 6 Conservative Substitutions = 0

X 10 20 30 40 50 60 70  
TCAGAGGAAGGCGTACGACCGCGCGCGCTGAGAGTGAACAAAGAGTGGTCTTGGCGGTG  
|||||  
TCAGAGGAAGGCGTACGACCGCGCGCGCTGAGAGTGAACAAAGAGTGGTCTTGGCGGTG  
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140  
CCACCCGCGCTGTGGTGCACGACCGCTACCGACGAGAGAGTGGTCCGACATCGGATACCTGTC  
|||||  
GCGAGC-----CATCAGCTCGGTCAGCGCTGACGAGAGAGAGTGGTCCGACATCGGATACCTGTC  
80 90 100 110 120 130 140

14. US-09-697-123B-5 Sequence 5, Application US/09697123B

Initial Score = 121 Optimized Score = 159 Significance = 0.09  
Residue Identity = 77% Matches = 166 Mismatches = 42  
Gaps = 6 Conservative Substitutions = 0

X 10 20 30 40 50 60 70  
TCAAGAGAAGGCGCTACGACCTGGCCCGCGCTGACAAAGTGAAGAGTGGCTTGGCGGTG  
|||||  
GCTTGCAGAGGCGCTACGACCTGGCCCGCGCTGACAAAGTGAAGAGTGGCTTGGCGGTG  
140 150 160 170 180 190 200 X

15. US-09-697-123B-7 (1-214)  
US-09-697-123B-2 Sequence 22, Application US/09697123B

Initial Score = 120 Optimized Score = 158 Significance = 0.07  
Residue Identity = 77% Matches = 165 Mismatches = 43  
Gaps = 6 Conservative Substitutions = 0

X 10 20 30 40 50 60 70  
TCAAGAGAAGGCGCTACGACCTGGCCCGCGCTGACAAAGTGAAGAGTGGCTTGGCGGTG  
|||||  
GCTTGCAGAGGCGCTACGACCTGGCCCGCGCTGACAAAGTGAAGAGTGGCTTGGCGGTG  
140 150 160 170 180 190 200 X

16. US-09-697-123B-7 (1-214)  
US-09-697-123B-2 Sequence 23, Application US/09697123B

Initial Score = 117 Optimized Score = 155 Significance = 0.00  
Residue Identity = 75% Matches = 162 Mismatches = 46  
Gaps = 6 Conservative Substitutions = 0

X 10 20 30 40 50 60 70  
TCAAGAGAAGGCGCTACGACCTGGCCCGCGCTGACAAAGTGAAGAGTGGCTTGGCGGTG  
|||||  
GCTTGCAGAGGCGCTACGACCTGGCCCGCGCTGACAAAGTGAAGAGTGGCTTGGCGGTG  
140 150 160 170 180 190 200 X

17. US-09-697-123B-7 (1-214)  
US-09-697-123B-1 Sequence 11, Application US/09697123B

Initial Score = 115 Optimized Score = 177 Significance = -0.05  
Residue Identity = 85% Matches = 184 Mismatches = 24  
Gaps = 6 Conservative Substitutions = 0

X 10 20 30 40 50 60 70  
TCAAGAGAAGGCGCTACGACCTGGCCCGCGCTGACAAAGTGAAGAGTGGCTTGGCGGTG  
|||||  
GCTTGCAGAGGCGCTACGACCTGGCCCGCGCTGACAAAGTGAAGAGTGGCTTGGCGGTG  
140 150 160 170 180 190 200 X

18. US-09-697-123B-7 (1-214)  
US-09-697-123B-2 Sequence 20, Application US/09697123B

Initial Score = 109 Optimized Score = 158 Significance = -0.19  
Residue Identity = 79% Matches = 170 Mismatches = 35  
Gaps = 9 Conservative Substitutions = 0

X 10 20 30 40 50 60 70  
TCAAGAGAAGGCGCTACGACCTGGCCCGCGCTGACAAAGTGAAGAGTGGCTTGGCGGTG  
|||||  
GCTTGCAGAGGCGCTACGACCTGGCCCGCGCTGACAAAGTGAAGAGTGGCTTGGCGGTG  
140 150 160 170 180 190 200 X

19. US-09-697-123B-7 (1-214)  
US-09-697-123B-1 Sequence 11, Application US/09697123B

Initial Score = 108 Optimized Score = 158 Significance = -0.21  
Residue Identity = 81% Matches = 175 Mismatches = 33  
Gaps = 6 Conservative Substitutions = 0

X 10 20 30 40 50 60 70  
TCAAGAGAAGGCGCTACGACCTGGCCCGCGCTGACAAAGTGAAGAGTGGCTTGGCGGTG  
|||||  
GCTTGCAGAGGCGCTACGACCTGGCCCGCGCTGACAAAGTGAAGAGTGGCTTGGCGGTG  
140 150 160 170 180 190 200 X

20. US-09-697-123B-7 (1-214)  
US-09-697-123B-2 Sequence 20, Application US/09697123B

Initial Score = 109 Optimized Score = 158 Significance = -0.19  
Residue Identity = 79% Matches = 170 Mismatches = 35  
Gaps = 9 Conservative Substitutions = 0

X 10 20 30 40 50 60 70  
TCAAGAGAAGGCGCTACGACCTGGCCCGCGCTGACAAAGTGAAGAGTGGCTTGGCGGTG  
|||||  
GCTTGCAGAGGCGCTACGACCTGGCCCGCGCTGACAAAGTGAAGAGTGGCTTGGCGGTG  
140 150 160 170 180 190 200 X

X 10 20 30 40 50 60 70  
TCAAGGAGAGGCGTACGACCTGCGCCCGCGGTGACAAAGTGAAGAAAGAGCTGGGCTTGGCGGTG  
TCAGAGGAGAGGCGTACGACCTGCGCCCGCGGTGACAAAGTGAAGAAAGAGCTGGGCTTGGCGGTG  
X 10 20 30 40 50 60 70  
CCAAACCCGCGCTGTGCTGACCTGCGCCCGCGGTGACAAAGTGAAGAAAGAGCTGGGCTTGGCGGTG  
GCCAGCC-----CATTCACAGCTGTGACCTGCGCCCGCGGTGACAAAGTGAAGAAAGAGCTGGGCTTGGCGGTG  
X 10 20 30 40 50 60 70  
150 160 170 180 190 200 210 X  
GCCCTGACGAGGCGGCGACGACGATGACCGCCCGCGGTGACAAAGTGAAGAAAGAGCTGGGCTTGGCGGTG  
GCTTGCACGAGGCGGCGACGATGACCGCCCGCGGTGACAAAGTGAAGAAAGAGCTGGGCTTGGCGGTG  
X 140 150 160 170 180 190 200 X

20. US-09-697-123B-7 (1-214)  
US-09-697-123B-1 Sequence 10, Application US/09697123B

Initial Score = 107 Optimized Score = 167 Significance = -0.24  
Residue Identity = 81% Matches = 174 Mismatches = 34  
Gaps = 6 Conservative Substitutions = 0

X 10 20 30 40 50 60 70  
TCAAGGAGAGGCGTACGACCTGCGCCCGCGGTGACAAAGTGAAGAAAGAGCTGGGCTTGGCGGTG  
TCAGAGGAGAGGCGTACGACCTGCGCCCGCGGTGACAAAGTGAAGAAAGAGCTGGGCTTGGCGGTG  
X 10 20 30 40 50 60 70  
80 90 100 110 120 130 140  
CCAAACCCGCGCTGTGCTGACCTGCGCCCGCGGTGACAAAGTGAAGAAAGAGCTGGGCTTGGCGGTG  
GCCAGCC-----CATTCACAGCTGTGACCTGCGCCCGCGGTGACAAAGTGAAGAAAGAGCTGGGCTTGGCGGTG  
X 150 160 170 180 190 200 X  
GCCCTGACGAGGCGGCGACGACGATGACCGCCCGCGGTGACAAAGTGAAGAAAGAGCTGGGCTTGGCGGTG  
GCTTGCACGAGGCGGCGACGATGACCGCCCGCGGTGACAAAGTGAAGAAAGAGCTGGGCTTGGCGGTG  
X 140 150 160 170 180 190 200 X

21. US-09-697-123B-7 (1-214)  
US-09-697-123B-1 Sequence 12, Application US/09697123B

Initial Score = 101 Optimized Score = 158 Significance = -0.38  
Residue Identity = 79% Matches = 170 Mismatches = 37  
Gaps = 7 Conservative Substitutions = 0

X 10 20 30 40 50 60 70  
TCAAGGAGAGGCGTACGACCTGCGCCCGCGGTGACAAAGTGAAGAAAGAGCTGGGCTTGGCGGTG  
TCAGAGGAGAGGCGTACGACCTGCGCCCGCGGTGACAAAGTGAAGAAAGAGCTGGGCTTGGCGGTG  
X 10 20 30 40 50 60 70  
80 90 100 110 120 130 140  
CCAAACCCGCGCTGTGCTGACCTGCGCCCGCGGTGACAAAGTGAAGAAAGAGCTGGGCTTGGCGGTG  
GCCAGCC-----CATTCACAGCTGTGACCTGCGCCCGCGGTGACAAAGTGAAGAAAGAGCTGGGCTTGGCGGTG  
X 150 160 170 180 190 200 X  
GCCCTGACGAGGCGGCGACGACGATGACCGCCCGCGGTGACAAAGTGAAGAAAGAGCTGGGCTTGGCGGTG  
GCTTGCACGAGGCGGCGACGATGACCGCCCGCGGTGACAAAGTGAAGAAAGAGCTGGGCTTGGCGGTG  
X 140 150 160 170 180 190 200 X

22. US-09-697-123B-7 (1-214)  
US-09-697-123B-4 Sequence 4, Application US/09697123B

Initial Score = 92 Optimized Score = 155 Significance = -0.59  
Residue Identity = 78% Matches = 167 Mismatches = 40  
Gaps = 7 Conservative Substitutions = 0

X 10 20 30 40 50 60 70  
TCAAGGAGAGGCGTACGACCTGCGCCCGCGGTGACAAAGTGAAGAAAGAGCTGGGCTTGGCGGTG  
TCAGAGGAGAGGCGTACGACCTGCGCCCGCGGTGACAAAGTGAAGAAAGAGCTGGGCTTGGCGGTG  
X 10 20 30 40 50 60 70  
80 90 100 110 120 130 140  
CCAAACCCGCGCTGTGCTGACCTGCGCCCGCGGTGACAAAGTGAAGAAAGAGCTGGGCTTGGCGGTG  
GCCAGCC-----CATTCACAGCTGTGACCTGCGCCCGCGGTGACAAAGTGAAGAAAGAGCTGGGCTTGGCGGTG  
X 150 160 170 180 190 200 X  
GCCCTGACGAGGCGGCGACGACGATGACCGCCCGCGGTGACAAAGTGAAGAAAGAGCTGGGCTTGGCGGTG  
GCTTGCACGAGGCGGCGACGATGACCGCCCGCGGTGACAAAGTGAAGAAAGAGCTGGGCTTGGCGGTG  
X 140 150 160 170 180 190 200 X

23. US-09-697-123B-7 (1-214)  
US-09-697-123B-1 Sequence 13, Application US/09697123B

Initial Score = 92 Optimized Score = 150 Significance = -0.59  
Residue Identity = 74% Matches = 171 Mismatches = 37  
Gaps = 21 Conservative Substitutions = 0

X 10 20 30 40 50 60 70  
TCAAGGAGAGGCGTACGACCTGCGCCCGCGGTGACAAAGTGAAGAAAGAGCTGGGCTTGGCGGTG  
TCAGAGGAGAGGCGTACGACCTGCGCCCGCGGTGACAAAGTGAAGAAAGAGCTGGGCTTGGCGGTG  
X 10 20 30 40 50 60 70  
80 90 100 110 120 130 140  
CCAAACCCGCGCTGTGCTGACCTGCGCCCGCGGTGACAAAGTGAAGAAAGAGCTGGGCTTGGCGGTG  
GCCAGCC-----CATTCACAGCTGTGACCTGCGCCCGCGGTGACAAAGTGAAGAAAGAGCTGGGCTTGGCGGTG  
X 150 160 170 180 190 200 X  
GCCCTGACGAGGCGGCGACGACGATGACCGCCCGCGGTGACAAAGTGAAGAAAGAGCTGGGCTTGGCGGTG  
GCTTGCACGAGGCGGCGACGATGACCGCCCGCGGTGACAAAGTGAAGAAAGAGCTGGGCTTGGCGGTG  
X 140 150 160 170 180 190 200 X

24. US-09-697-123B-7 (1-214)  
US-09-697-123B-9 Sequence 9, Application US/09697123B

Initial Score = 91 Optimized Score = 154 Significance = -0.62  
Residue Identity = 74% Matches = 170 Mismatches = 38  
Gaps = 21 Conservative Substitutions = 0

X 10 20 30 40 50 60 70  
TCAAGGAGAGGCGTACGACCTGCGCCCGCGGTGACAAAGTGAAGAAAGAGCTGGGCTTGGCGGTG  
TCAGAGGAGAGGCGTACGACCTGCGCCCGCGGTGACAAAGTGAAGAAAGAGCTGGGCTTGGCGGTG  
X 10 20 30 40 50 60 70  
80 90 100 110 120 130 140  
CCAAACCCGCGCTGTGCTGACCTGCGCCCGCGGTGACAAAGTGAAGAAAGAGCTGGGCTTGGCGGTG  
GCCAGCC-----CATTCACAGCTGTGACCTGCGCCCGCGGTGACAAAGTGAAGAAAGAGCTGGGCTTGGCGGTG  
X 150 160 170 180 190 200 X  
GCCCTGACGAGGCGGCGACGACGATGACCGCCCGCGGTGACAAAGTGAAGAAAGAGCTGGGCTTGGCGGTG  
GCTTGCACGAGGCGGCGACGATGACCGCCCGCGGTGACAAAGTGAAGAAAGAGCTGGGCTTGGCGGTG  
X 140 150 160 170 180 190 200 X

	80	90	100	110	120	130	
150	160			170	180	190	200
GCCTGCACGAGCGCCAGA	-----	-----	CCAAGATGACCGCGCGGCGCTGAGTCCCGGTCG				
140	150	160	170	180	190	200	210
GCCTGCACGAGCGCCCGTACGGATGCGCCAGCGCCGCTCATGACTGTCCCGGCGCATCGAGGTGCGGATGg							
210	X						
AGTCGACGACAT							
AGACGACGACAT							
220	X						

> 0 <  
0100 Intelligenetics  
> 0 <  
FastDB - Fast Pairwise Comparison of Sequences  
Release 5.4

Results file us-09-697-123b-6.res made by shanley on Wed 13 Nov 102 14:04:08-PSR.

Query sequence being compared: US-09-697-123B-6 (1-208)  
Number of sequences searched: 26  
Number of scores above cutoff: 25

Results of the initial comparison of US-09-697-123B-6 (1-208) with:  
File: US09697123B.seq

```

100-
N -
U -
B -
M -
E -
R -
O -
F -
S -
E -
U -
E -
C -
S -
SCORE 0 231 46 69 92 116 139 162 185 208
STDEV -2 -1 0 1

```

## PARAMETERS

Similarity matrix Unitary 1 K-tuple 4  
Mismatch penalty 1 Joining penalty 30  
Gap penalty 5.00 Window size 207  
Gap size penalty 0.33  
Cutoff score 12  
Randomization group 0

## SEARCH STATISTICS

Scores: Mean 155 Median 179 Standard Deviation 51.93  
Times: CPU 00:00:00.00 Total Elapsed 00:00:00.00  
Number of residues: 5077  
Number of sequences searched: 26  
Number of scores above cutoff: 25

The scores below are sorted by initial score.  
Significance is calculated based on initial score.

3 100% identical sequences to the query sequence were found:

Sequence Name	Description	Length	Int. Opt. Score	Sig. Frame
---------------	-------------	--------	-----------------	------------

1. US-09-697-123B-1 Sequence 16, Application 208 208 208 1.02 0  
2. US-09-697-123B-6 Sequence 6, Application U 208 208 208 1.02 0  
3. US-09-697-123B-5 Sequence 5, Application U 208 208 208 1.02 0  
The list of other best scores is:

Sequence Name	Description	Length	Int. Opt. Score	Sig. Frame	
4. US-09-697-123B-2	Sequence 22, Application **** 1 standard deviation above mean ****	208	207	207	1.00 0
5. US-09-697-123B-1	Sequence 10, Application **** 0 standard deviation from mean ****	208	186	186	0.60 0
6. US-09-697-123B-1	Sequence 11, Application	208	185	185	0.58 0
7. US-09-697-123B-3	Sequence 3, Application U	208	185	185	0.58 0
8. US-09-697-123B-1	Sequence 15, Application	208	184	184	0.56 0
9. US-09-697-123B-2	Sequence 23, Application	208	182	182	0.52 0
10. US-09-697-123B-1	Sequence 19, Application	208	182	182	0.52 0
11. US-09-697-123B-8	Sequence 8, Application U	208	182	182	0.52 0
12. US-09-697-123B-1	Sequence 1, Application U	208	182	182	0.52 0
13. US-09-697-123B-2	Sequence 2, Application U	208	181	181	0.50 0
14. US-09-697-123B-1	Sequence 17, Application	208	178	178	0.44 0
15. US-09-697-123B-2	Sequence 24, Application	208	172	172	0.33 0
16. US-09-697-123B-2	Sequence 20, Application	205	161	171	0.12 0
17. US-09-697-123B-1	Sequence 12, Application	207	157	176	0.04 0
18. US-09-697-123B-4	Sequence 4, Application U	207	147	178	-0.15 0
19. US-09-697-123B-9	Sequence 9, Application U	223	144	170	-0.21 0
20. US-09-697-123B-1	Sequence 13, Application	223	137	162	-0.35 0
21. US-09-697-123B-7	Sequence 7, Application U	214	121	159	-0.65 0
22. US-09-697-123B-1	Sequence 18, Application	211	113	159	-0.81 0
23. US-09-697-123B-1	Sequence 14, Application	214	110	168	-0.87 0
24. US-09-697-123B-2	Sequence 21, Application	214	109	162	-0.89 0
25. US-09-697-123B-2	Sequence 25, Application **** 2 standard deviations below mean ****	19	19	19	-2.62 0

## 1. US-09-697-123B-6 (1-208)

US-09-697-123B-1 Sequence 16, Application US/09697123B

Initial Score - 208 Optimized Score - 208 Significance - 1.02  
Residue Identity - 100% Matches - 208 Mismatches - 0  
Gaps - 0 Conservative Substitutions - 0

```

X 10 20 30 40 50 60 70
TCAAGGAGAGCGCTACGACCTGCGCCGCTGCGCTATTAAGTCAACAAGAGCTGCGGCTGCATGCTG
|||||
TCAAGGAGAGCGCTACGACCTGCGCCGCTGCGCTATTAAGTCAACAAGAGCTGCGGCTGCATGCTG
X 10 20 30 40 50 60 70

```

```

80 90 100 110 120 130 140
GCGAGCCCATCAGCTGCTGCGCTGACCGAAGAGAGCTGCGCTATTAAGTCAACAAGAGCTGCGGCTGC
|||||
GCGAGCCCATCAGCTGCTGCGCTGACCGAAGAGAGCTGCGCTATTAAGTCAACAAGAGCTGCGGCTGC
80 90 100 110 120 130 140

```

```

150 160 170 180 190 200
ACGAGGTCAGACCAAGATGACCTGCGCGCGCTGAGGTGCGGTGAACCGACGACAT
|||||
ACGAGGTCAGACCAAGATGACCTGCGCGCGCTGAGGTGCGGTGAACCGACGACAT
X 150 160 170 180 190 200

```

## 2. US-09-697-123B-6 (1-208)

US-09-697-123B-6 Sequence 6, Application US/09697123B

Initial Score - 208 Optimized Score - 208 Significance - 1.02  
Residue Identity - 100% Matches - 208 Mismatches - 0  
Gaps - 0 Conservative Substitutions - 0

```

X 10 20 30 40 50 60 70
TCAAGGAGAGCGCTACGACCTGCGCGCTGCGCTATTAAGTCAACAAGAGCTGCGGCTGCATGCTG
|||||
TCAAGGAGAGCGCTACGACCTGCGCGCTGCGCTATTAAGTCAACAAGAGCTGCGGCTGCATGCTG
X 10 20 30 40 50 60 70

```

X 10 20 30 40 50 60 70

CGGAGCCCATACGTCGTCGACGCTGACCGAAGAAGACGTGTGGCCCATCGAATATCTGGTCCGCTTGC  
|||||  
GCGAGCCCATACGTCGTCGACGCTGACCGAAGAAGACGTGTGGCCCATCGAATATCTGGTCCGCTTGC  
80 90 100 110 120 130 140  
GCGAGCCCATACGTCGTCGACGCTGACCGAAGAAGACGTGTGGCCCATCGAATATCTGGTCCGCTTGC  
80 90 100 110 120 130 140

ACGAGGCTCAGACACCGATGACGCTTCCGGGGCGGCTCGAGGTGCCCGGTGGAACCGACGACAT X  
|||||  
ACGAGGCTCAGACACCGATGACGCTTCCGGGGCGGCTCGAGGTGCCCGGTGGAACCGACGACAT  
150 160 170 180 190 200  
ACGAGGCTCAGACACCGATGACGCTTCCGGGGCGGCTCGAGGTGCCCGGTGGAACCGACGACAT X

## 3. US-09-697-123B-6 (1-208)

US-09-697-123B-5 Sequence 5, Application US/09697123B

Initial Score = 208 Optimized Score = 208 Significance = 1.02  
Residue Identity = 100% Matches = 208 Mismatches = 0  
Gaps = 0 Conservative Substitutions = 0

X 10 20 30 40 50 60 70

TCAGAGAGAACCGCTACGACCTGGCCCGCTCGCTGCTATAGGTCAACAGAAGCTCGGCTGCATGTGC  
|||||  
TCAGAGAGAACCGCTACGACCTGGCCCGCTCGCTGCTATAGGTCAACAGAAGCTCGGCTGCATGTGC  
10 20 30 40 50 60 70

CGGAGCCCATACGTCGTCGACGCTGACCGAAGAAGACGTGTGGCCCATCGAATATCTGGTCCGCTTGC  
|||||  
CGGAGCCCATACGTCGTCGACGCTGACCGAAGAAGACGTGTGGCCCATCGAATATCTGGTCCGCTTGC  
80 90 100 110 120 130 140  
CGGAGCCCATACGTCGTCGACGCTGACCGAAGAAGACGTGTGGCCCATCGAATATCTGGTCCGCTTGC  
80 90 100 110 120 130 140

ACGAGGCTCAGACACCGATGACGCTTCCGGGGCGGCTCGAGGTGCCCGGTGGAACCGACGACAT X  
|||||  
ACGAGGCTCAGACACCGATGACGCTTCCGGGGCGGCTCGAGGTGCCCGGTGGAACCGACGACAT  
150 160 170 180 190 200  
ACGAGGCTCAGACACCGATGACGCTTCCGGGGCGGCTCGAGGTGCCCGGTGGAACCGACGACAT X

## 4. US-09-697-123B-6 (1-208)

US-09-697-123B-2 Sequence 22, Application US/09697123B

Initial Score = 207 Optimized Score = 207 Significance = 1.00  
Residue Identity = 99% Matches = 207 Mismatches = 1  
Gaps = 0 Conservative Substitutions = 0

X 10 20 30 40 50 60 70

TCAGAGAGAACCGCTACGACCTGGCCCGCTCGCTGCTATAGGTCAACAGAAGCTCGGCTGCATGTGC  
|||||  
TCAGAGAGAACCGCTACGACCTGGCCCGCTCGCTGCTATAGGTCAACAGAAGCTCGGCTGCATGTGC  
10 20 30 40 50 60 70

CGGAGCCCATACGTCGTCGACGCTGACCGAAGAAGACGTGTGGCCCATCGAATATCTGGTCCGCTTGC  
|||||  
CGGAGCCCATACGTCGTCGACGCTGACCGAAGAAGACGTGTGGCCCATCGAATATCTGGTCCGCTTGC  
80 90 100 110 120 130 140  
CGGAGCCCATACGTCGTCGACGCTGACCGAAGAAGACGTGTGGCCCATCGAATATCTGGTCCGCTTGC  
80 90 100 110 120 130 140

ACGAGGCTCAGACACCGATGACGCTTCCGGGGCGGCTCGAGGTGCCCGGTGGAACCGACGACAT X  
|||||  
ACGAGGCTCAGACACCGATGACGCTTCCGGGGCGGCTCGAGGTGCCCGGTGGAACCGACGACAT  
150 160 170 180 190 200  
ACGAGGCTCAGACACCGATGACGCTTCCGGGGCGGCTCGAGGTGCCCGGTGGAACCGACGACAT X

150 160 170 180 190 200 X  
ACGAGGCTCAGACACCGATGACGCTTCCGGGGCGGCTCGAGGTGCCCGGTGGAACCGACGACAT  
|||||  
ACGAGGCTCAGACACCGATGACGCTTCCGGGGCGGCTCGAGGTGCCCGGTGGAACCGACGACAT  
150 160 170 180 190 200  
ACGAGGCTCAGACACCGATGACGCTTCCGGGGCGGCTCGAGGTGCCCGGTGGAACCGACGACAT X

## 5. US-09-697-123B-6 (1-208)

US-09-697-123B-1 Sequence 10, Application US/09697123B

Initial Score = 186 Optimized Score = 186 Significance = 0.60  
Residue Identity = 89% Matches = 186 Mismatches = 22  
Gaps = 0 Conservative Substitutions = 0

X 10 20 30 40 50 60 70

TCAGAGAGAACCGCTACGACCTGGCCCGCTCGCTGCTATAGGTCAACAGAAGCTCGGCTGCATGTGC  
|||||  
TCAGAGAGAACCGCTACGACCTGGCCCGCTCGCTGCTATAGGTCAACAGAAGCTCGGCTGCATGTGC  
10 20 30 40 50 60 70  
TCAGAGAGAACCGCTACGACCTGGCCCGCTCGCTGCTATAGGTCAACAGAAGCTCGGCTGCATGTGC  
10 20 30 40 50 60 70

CGGAGCCCATACGTCGTCGACGCTGACCGAAGAAGACGTGTGGCCCATCGAATATCTGGTCCGCTTGC  
|||||  
CGGAGCCCATACGTCGTCGACGCTGACCGAAGAAGACGTGTGGCCCATCGAATATCTGGTCCGCTTGC  
80 90 100 110 120 130 140  
CGGAGCCCATACGTCGTCGACGCTGACCGAAGAAGACGTGTGGCCCATCGAATATCTGGTCCGCTTGC  
80 90 100 110 120 130 140

ACGAGGCTCAGACACCGATGACGCTTCCGGGGCGGCTCGAGGTGCCCGGTGGAACCGACGACAT X  
|||||  
ACGAGGCTCAGACACCGATGACGCTTCCGGGGCGGCTCGAGGTGCCCGGTGGAACCGACGACAT  
150 160 170 180 190 200  
ACGAGGCTCAGACACCGATGACGCTTCCGGGGCGGCTCGAGGTGCCCGGTGGAACCGACGACAT X

## 6. US-09-697-123B-6 (1-208)

US-09-697-123B-1 Sequence 11, Application US/09697123B

Initial Score = 185 Optimized Score = 185 Significance = 0.58  
Residue Identity = 88% Matches = 185 Mismatches = 23  
Gaps = 0 Conservative Substitutions = 0

X 10 20 30 40 50 60 70

TCAGAGAGAACCGCTACGACCTGGCCCGCTCGCTGCTATAGGTCAACAGAAGCTCGGCTGCATGTGC  
|||||  
TCAGAGAGAACCGCTACGACCTGGCCCGCTCGCTGCTATAGGTCAACAGAAGCTCGGCTGCATGTGC  
10 20 30 40 50 60 70

CGGAGCCCATACGTCGTCGACGCTGACCGAAGAAGACGTGTGGCCCATCGAATATCTGGTCCGCTTGC  
|||||  
CGGAGCCCATACGTCGTCGACGCTGACCGAAGAAGACGTGTGGCCCATCGAATATCTGGTCCGCTTGC  
80 90 100 110 120 130 140  
CGGAGCCCATACGTCGTCGACGCTGACCGAAGAAGACGTGTGGCCCATCGAATATCTGGTCCGCTTGC  
80 90 100 110 120 130 140

ACGAGGCTCAGACACCGATGACGCTTCCGGGGCGGCTCGAGGTGCCCGGTGGAACCGACGACAT X  
|||||  
ACGAGGCTCAGACACCGATGACGCTTCCGGGGCGGCTCGAGGTGCCCGGTGGAACCGACGACAT  
150 160 170 180 190 200  
ACGAGGCTCAGACACCGATGACGCTTCCGGGGCGGCTCGAGGTGCCCGGTGGAACCGACGACAT X

## 7. US-09-697-123B-6 (1-208)

US-09-697-123B-3 Sequence 3, Application US/09697123B

Initial Score = 185 Optimized Score = 185 Significance = 0.58  
Residue Identity = 88% Matches = 185 Mismatches = 23  
Gaps = 0 Conservative Substitutions = 0

X 10 20 30 40 50 60 70

TCAGAGAGAACCGCTACGACCTGGCCCGCTCGCTGCTATAGGTCAACAGAAGCTCGGCTGCATGTGC  
|||||  
TCAGAGAGAACCGCTACGACCTGGCCCGCTCGCTGCTATAGGTCAACAGAAGCTCGGCTGCATGTGC  
10 20 30 40 50 60 70

CGGAGCCCATACGTCGTCGACGCTGACCGAAGAAGACGTGTGGCCCATCGAATATCTGGTCCGCTTGC  
|||||  
CGGAGCCCATACGTCGTCGACGCTGACCGAAGAAGACGTGTGGCCCATCGAATATCTGGTCCGCTTGC  
80 90 100 110 120 130 140  
CGGAGCCCATACGTCGTCGACGCTGACCGAAGAAGACGTGTGGCCCATCGAATATCTGGTCCGCTTGC  
80 90 100 110 120 130 140

ACGAGGCTCAGACACCGATGACGCTTCCGGGGCGGCTCGAGGTGCCCGGTGGAACCGACGACAT X  
|||||  
ACGAGGCTCAGACACCGATGACGCTTCCGGGGCGGCTCGAGGTGCCCGGTGGAACCGACGACAT  
150 160 170 180 190 200  
ACGAGGCTCAGACACCGATGACGCTTCCGGGGCGGCTCGAGGTGCCCGGTGGAACCGACGACAT X

## 8. US-09-697-123B-6 (1-208)

## US-09-697-123B-1 Sequence 15, Application US/09697123B

Initial Score = 184 Optimized Score = 184 Significance = 0.56  
Residue Identity = 88% Matches = 184 Mismatches = 24  
Gaps = 0 Conservative Substitutions = 0

X  
TCGAAGAGAAAGCGCTACGACCTGCGCCCGCGTGCCTATTAAGGTCAACAGAAAGCTGGCGCTGCATGTCG  
TCAGAGAGAAAGCGCTACGACCTGCGCCCGCGTGCCTATTAAGGTCAACAGAAAGCTGGCGCTGCATGTCG  
X  
10 20 30 40 50 60 70

80 90 100 110 120 130 140  
GCGAGCCCATCAGCTGCTGACGCTGACCGGAGAGAGAGCTGCTGCGCATCAATCAATATCTGCTCCGCTTCG  
GTAAGCCGATCAGCAGCTGCTGACCGGAGAGAGAGCTGCTGCGCATCAATCAATATCTGCTCCGCTTCG  
X  
150 160 170 180 190 200  
ACGAGGCTCAGACCCAGAGATGACCTGCTCCCGCGCGCATCAGAGTCCCGGTGGAGACCGACGACAT X

150 160 170 180 190 200  
ACGAGGCTCAGACCCAGAGATGACCTGCTCCCGCGCGCATCAGAGTCCCGGTGGAGACCGACGACAT X  
ACGAGGCTCAGACCCAGAGATGACCTGCTCCCGCGCGCATCAGAGTCCCGGTGGAGACCGACGACAT X  
X  
150 160 170 180 190 200

9. US-09-697-123B-6 (1-208)  
US-09-697-123B-2 Sequence 23, Application US/09697123B

Initial Score = 182 Optimized Score = 182 Significance = 0.52  
Residue Identity = 87% Matches = 182 Mismatches = 26  
Gaps = 0 Conservative Substitutions = 0

X  
TCGAAGAGAAAGCGCTACGACCTGCGCCCGCGTGCCTATTAAGGTCAACAGAAAGCTGGCGCTGCATGTCG  
TCAGAGAGAAAGCGCTACGACCTGCGCCCGCGTGCCTATTAAGGTCAACAGAAAGCTGGCGCTGCATGTCG  
X  
10 20 30 40 50 60 70

80 90 100 110 120 130 140  
GCGAGCCCATCAGCTGCTGACGCTGACCGGAGAGAGAGCTGCTGCGCATCAATCAATATCTGCTCCGCTTCG  
GTAAGCCGATCAGCAGCTGCTGACCGGAGAGAGAGCTGCTGCGCATCAATCAATATCTGCTCCGCTTCG  
X  
150 160 170 180 190 200  
ACGAGGCTCAGACCCAGAGATGACCTGCTCCCGCGCGCATCAGAGTCCCGGTGGAGACCGACGACAT X

150 160 170 180 190 200  
ACGAGGCTCAGACCCAGAGATGACCTGCTCCCGCGCGCATCAGAGTCCCGGTGGAGACCGACGACAT X  
ACGAGGCTCAGACCCAGAGATGACCTGCTCCCGCGCGCATCAGAGTCCCGGTGGAGACCGACGACAT X  
X  
150 160 170 180 190 200

10. US-09-697-123B-6 (1-208)  
US-09-697-123B-1 Sequence 19, Application US/09697123B

Initial Score = 182 Optimized Score = 182 Significance = 0.52  
Residue Identity = 87% Matches = 182 Mismatches = 26  
Gaps = 0 Conservative Substitutions = 0

X  
TCGAAGAGAAAGCGCTACGACCTGCGCCCGCGTGCCTATTAAGGTCAACAGAAAGCTGGCGCTGCATGTCG  
TCAGAGAGAAAGCGCTACGACCTGCGCCCGCGTGCCTATTAAGGTCAACAGAAAGCTGGCGCTGCATGTCG  
X  
10 20 30 40 50 60 70

80 90 100 110 120 130 140  
GCGAGCCCATCAGCTGCTGACGCTGACCGGAGAGAGAGCTGCTGCGCATCAATCAATATCTGCTCCGCTTCG  
GTAAGCCGATCAGCAGCTGCTGACCGGAGAGAGAGCTGCTGCGCATCAATCAATATCTGCTCCGCTTCG  
X  
150 160 170 180 190 200  
ACGAGGCTCAGACCCAGAGATGACCTGCTCCCGCGCGCATCAGAGTCCCGGTGGAGACCGACGACAT X

ACGAGGCCAGACCCAGATGACCTGCTCCCGCGCGCTGAGAGTCCCGGTGGAGAGACGACAT X  
150 160 170 180 190 200

11. US-09-697-123B-8 Sequence 8, Application US/09697123B  
US-09-697-123B-1 Sequence 1, Application US/09697123B

Initial Score = 182 Optimized Score = 182 Significance = 0.52  
Residue Identity = 87% Matches = 182 Mismatches = 26  
Gaps = 0 Conservative Substitutions = 0

X  
TCGAAGAGAAAGCGCTACGACCTGCGCCCGCGTGCCTATTAAGGTCAACAGAAAGCTGGCGCTGCATGTCG  
TCAGAGAGAAAGCGCTACGACCTGCGCCCGCGTGCCTATTAAGGTCAACAGAAAGCTGGCGCTGCATGTCG  
X  
10 20 30 40 50 60 70

80 90 100 110 120 130 140  
GCGAGCCCATCAGCTGCTGACGCTGACCGGAGAGAGAGCTGCTGCGCATCAATCAATATCTGCTCCGCTTCG  
GTAAGCCGATCAGCAGCTGCTGACCGGAGAGAGAGCTGCTGCGCATCAATCAATATCTGCTCCGCTTCG  
X  
150 160 170 180 190 200  
ACGAGGCTCAGACCCAGAGATGACCTGCTCCCGCGCGCATCAGAGTCCCGGTGGAGACCGACGACAT X

150 160 170 180 190 200  
ACGAGGCTCAGACCCAGAGATGACCTGCTCCCGCGCGCATCAGAGTCCCGGTGGAGACCGACGACAT X  
ACGAGGCTCAGACCCAGAGATGACCTGCTCCCGCGCGCATCAGAGTCCCGGTGGAGACCGACGACAT X  
X  
150 160 170 180 190 200

12. US-09-697-123B-6 (1-208)  
US-09-697-123B-1 Sequence 1, Application US/09697123B

Initial Score = 182 Optimized Score = 182 Significance = 0.52  
Residue Identity = 87% Matches = 182 Mismatches = 26  
Gaps = 0 Conservative Substitutions = 0

X  
TCGAAGAGAAAGCGCTACGACCTGCGCCCGCGTGCCTATTAAGGTCAACAGAAAGCTGGCGCTGCATGTCG  
TCAGAGAGAAAGCGCTACGACCTGCGCCCGCGTGCCTATTAAGGTCAACAGAAAGCTGGCGCTGCATGTCG  
X  
10 20 30 40 50 60 70

80 90 100 110 120 130 140  
GCGAGCCCATCAGCTGCTGACGCTGACCGGAGAGAGAGCTGCTGCGCATCAATCAATATCTGCTCCGCTTCG  
GTAAGCCGATCAGCAGCTGCTGACCGGAGAGAGAGCTGCTGCGCATCAATCAATATCTGCTCCGCTTCG  
X  
150 160 170 180 190 200  
ACGAGGCTCAGACCCAGAGATGACCTGCTCCCGCGCGCATCAGAGTCCCGGTGGAGACCGACGACAT X

150 160 170 180 190 200  
ACGAGGCTCAGACCCAGAGATGACCTGCTCCCGCGCGCATCAGAGTCCCGGTGGAGACCGACGACAT X  
ACGAGGCTCAGACCCAGAGATGACCTGCTCCCGCGCGCATCAGAGTCCCGGTGGAGACCGACGACAT X  
X  
150 160 170 180 190 200

13. US-09-697-123B-6 (1-208)  
US-09-697-123B-2 Sequence 2, Application US/09697123B

Initial Score = 181 Optimized Score = 181 Significance = 0.50  
Residue Identity = 87% Matches = 181 Mismatches = 27  
Gaps = 0 Conservative Substitutions = 0

X  
TCGAAGAGAAAGCGCTACGACCTGCGCCCGCGTGCCTATTAAGGTCAACAGAAAGCTGGCGCTGCATGTCG  
TCAGAGAGAAAGCGCTACGACCTGCGCCCGCGTGCCTATTAAGGTCAACAGAAAGCTGGCGCTGCATGTCG  
X  
10 20 30 40 50 60 70

80 90 100 110 120 130 140  
GCGAGCCCATCAGCTGCTGACGCTGACCGGAGAGAGAGCTGCTGCGCATCAATCAATATCTGCTCCGCTTCG  
GTAAGCCGATCAGCAGCTGCTGACCGGAGAGAGAGCTGCTGCGCATCAATCAATATCTGCTCCGCTTCG  
X  
150 160 170 180 190 200  
ACGAGGCTCAGACCCAGAGATGACCTGCTCCCGCGCGCATCAGAGTCCCGGTGGAGACCGACGACAT X

Initial Score	-	144	Optimized Score	-	170	Significance	-	-0.21
Residue Identity	-	80%	Matches	-	180	Mismatches	-	28
Gaps	-	15	Conservative Substitutions	-			-	0



X  
 10  
 20  
 30  
 40  
 50  
 60  
 70  
 TCAAGAGAAGCGCTACGACCTCGGCCGCGCTGGCTTAAGTCAACGAAGCTCGGCTGCATATGCG  
 TCAAGAGAAGCGCTACGACCTCGGCCGCGCTGGCTTAAGTCAACGAAGCTCGGCTGCATATGCG  
 X  
 10  
 20  
 30  
 40  
 50  
 60  
 70  
 GCGAGCC-----CATCACGTTGTTGACGCTTACCGAAGAAGACGTCGTGTCGCCACCATCGAATATCTGTTC

X  
10 20 30 40 50 60 70  
TCAAGGAGAGCGCTACGACCTGGCCCGCGTGCCTATTAAGGTCAACAAGAGCTCGGGTGCATGCG  
TCAAGAGAGAGCGCTACGATCTGGCCCGCGTGGCTGCTACAAAGTGAACAAGAGTGGCCCTGGCGGCA



> 0 <  
 01 10 Intelligenetics  
 > 0 <  
 FastDB - Fast Pairwise Comparison of Sequences  
 Release 5.4

Results file us-09-697-123b-5.res made by shanley on Wed 13 Nov 102 14:03:13-PST.

Query sequence being compared: US-09-697-123B-5 (1-208)  
 Number of sequences searched: 26  
 Number of scores above cutoff: 25

Results of the initial comparison of US-09-697-123B-5 (1-208) with:  
 File : US09697123B.seq

```

100-
N -
U -
M -
B -
E -
R -
O -
F 10-
S -
E 5-
O -
U -
E -
N -
C -
S 0-
SCORE 0 231 46 69 92 116 139 162 185 208
STDV -2 -1 -1 0 0 1 1

```

## PARAMETERS

Similarity matrix Unitary 1 K-tuple 4  
 Mismatch penalty 1 Joining penalty 30  
 Gap penalty 5.00 Window size 207  
 Gap size penalty 0.33  
 Cutoff score 12  
 Randomization group 0

## SEARCH STATISTICS

Scores: Mean 155 Median 179 Standard Deviation 51.93  
 Times: CPU 00:00:00.00 Total Elapsed 00:00:00.00  
 Number of residues: 5077  
 Number of sequences searched: 26  
 Number of scores above cutoff: 25

The scores below are sorted by initial score.  
 Significance is calculated based on initial score.

3 100% identical sequences to the query sequence were found:

Sequence Name	Description	Length	Int. Opt. Score	Sig. Frame
---------------	-------------	--------	-----------------	------------

1. US-09-697-123B-1 Sequence 16, Application U 208 208 208 1.02 0  
 2. US-09-697-123B-6 Sequence 6, Application U 208 208 208 1.02 0  
 3. US-09-697-123B-5 Sequence 5, Application U 208 208 208 1.02 0  
 The list of other best scores is:

Sequence Name	Description	Length	Int. Opt. Score	Sig. Frame
4. US-09-697-123B-2 Sequence 22, Application U	**** 1 standard deviation above mean ****	208	207	1.00 0
5. US-09-697-123B-1 Sequence 10, Application U	**** 0 standard deviation from mean ****	208	186	0.60 0
6. US-09-697-123B-1 Sequence 11, Application U		208	185	0.58 0
7. US-09-697-123B-3 Sequence 3, Application U		208	185	0.58 0
8. US-09-697-123B-1 Sequence 15, Application U		208	184	0.56 0
9. US-09-697-123B-2 Sequence 23, Application U		208	182	0.52 0
10. US-09-697-123B-1 Sequence 19, Application U		208	182	0.52 0
11. US-09-697-123B-8 Sequence 8, Application U		208	182	0.52 0
12. US-09-697-123B-1 Sequence 1, Application U		208	182	0.52 0
13. US-09-697-123B-2 Sequence 2, Application U		208	181	0.50 0
14. US-09-697-123B-1 Sequence 17, Application U		208	178	0.44 0
15. US-09-697-123B-2 Sequence 24, Application U		208	172	0.33 0
16. US-09-697-123B-2 Sequence 12, Application U		205	161	0.12 0
17. US-09-697-123B-1 Sequence 4, Application U		207	157	0.04 0
18. US-09-697-123B-4 Sequence 9, Application U		207	147	0.15 0
19. US-09-697-123B-9 Sequence 7, Application U		223	144	0.21 0
20. US-09-697-123B-1 Sequence 13, Application U		223	137	0.35 0
21. US-09-697-123B-7 Sequence 7, Application U		214	121	0.62 0
22. US-09-697-123B-1 Sequence 18, Application U		211	113	0.59 0
23. US-09-697-123B-1 Sequence 14, Application U		214	110	0.87 0
24. US-09-697-123B-2 Sequence 21, Application U		214	109	0.89 0
25. US-09-697-123B-2 Sequence 25, Application U	**** 2 standard deviations below mean ****	19	19	-2.62 0

## 1. US-09-697-123B-5 (1-208)

US-09-697-123B-1 Sequence 16, Application US/09697123B

Initial Score = 208 Optimized Score = 208 Significance = 1.02  
 Residue Identity = 100% Matches = 208 Mismatches = 0  
 Gaps = 0 Conservative Substitutions = 0

X 10 20 30 40 50 60 70  
 TCAAGGAGAGGCGCTACGACCTGCGCCGCGCTGCTATTAAGTCAACAAGAGCTGGCGTCAATGTCG  
 TCAAGGAGAGGCGCTACGACCTGCGCCGCGCTGCTATTAAGTCAACAAGAGCTGGCGTCAATGTCG  
 X 10 20 30 40 50 60 70

CGGAGCCCATCAAGCTGCTGACGCTGACCGAAGAGAGCTGCGCCGCGCTGCTATTAAGTCAACAAGAGCTGGCGTCAATGTCG  
 CGGAGCCCATCAAGCTGCTGACGCTGACCGAAGAGAGCTGCGCCGCGCTGCTATTAAGTCAACAAGAGCTGGCGTCAATGTCG  
 X 80 90 100 110 120 130 140  
 GCGAGCCCATCAAGCTGCTGACGCTGACCGAAGAGAGCTGCGCCGCGCTGCTATTAAGTCAACAAGAGCTGGCGTCAATGTCG  
 GCGAGCCCATCAAGCTGCTGACGCTGACCGAAGAGAGCTGCGCCGCGCTGCTATTAAGTCAACAAGAGCTGGCGTCAATGTCG  
 X 80 90 100 110 120 130 140

150 160 170 180 190 200 X  
 ACGAGGCTGACGACGATGACCGCTTCCGGCGCGCTGACGCTGCGCCGCGCTGCTATTAAGTCAACAAGAGCTGGCGTCAATGTCG  
 ACGAGGCTGACGACGATGACCGCTTCCGGCGCGCTGACGCTGCGCCGCGCTGCTATTAAGTCAACAAGAGCTGGCGTCAATGTCG  
 X 150 160 170 180 190 200 X

## 2. US-09-697-123B-5 (1-208)

US-09-697-123B-6 Sequence 6, Application US/09697123B

Initial Score = 208 Optimized Score = 208 Significance = 1.02  
 Residue Identity = 100% Matches = 208 Mismatches = 0  
 Gaps = 0 Conservative Substitutions = 0

X 10 20 30 40 50 60 70  
 TCAAGGAGAGGCGCTACGACCTGCGCCGCGCTGCTATTAAGTCAACAAGAGCTGGCGTCAATGTCG  
 TCAAGGAGAGGCGCTACGACCTGCGCCGCGCTGCTATTAAGTCAACAAGAGCTGGCGTCAATGTCG  
 X 10 20 30 40 50 60 70

X 10 20 30 40 50 60 70

GGAGCCCATACGTCGTCGACGCTGACCCAGAGAGAGCTGCGCCACCATCGAATATCTGTGCGCTTGC  
|||||  
GCGAGCCCATACGTCGTCGACGCTGACCCAGAGAGAGCTGCGCCACCATCGAATATCTGTGCGCTTGC  
80 90 100 110 120 130 140

ACGAGGGTCAGACACCATGACGCTTCCGGGGCGGCTCGAGGTGCGCGGTGGAACCGAGACAT X  
|||||  
ACGAGGGTCAGACACCATGACGCTTCCGGGGCGGCTCGAGGTGCGCGGTGGAACCGAGACAT  
150 160 170 180 190 200

3. US-09-697-123B-5 (1-208)

US-09-697-123B-5 Sequence 5, Application US/09697123B

Initial Score = 208 Optimized Score = 208 Significance = 1.02  
Residue Identity = 100% Matches = 208 Mismatches = 0  
Gaps = 0 Conservative Substitutions = 0

X 10 20 30 40 50 60 70  
TCAAGAGAAAGCCGTACGACCTGCGCCCGCGCTGCTGCTATTAAGCTCAACAGAGAGCTGCGGCTGATGTG  
|||||  
TCAAGAGAAAGCCGTACGACCTGCGCCCGCGCTGCTGCTATTAAGCTCAACAGAGAGCTGCGGCTGATGTG  
10 20 30 40 50 60 70

GGGAGCCCATACGTCGTCGACGCTTACCCGAGAGAGAGCTGCGCTGATTAAGCTCAACAGAGAGCTGCGGCTGATGTG  
|||||  
GGGAGCCCATACGTCGTCGACGCTTACCCGAGAGAGAGCTGCGCTGATTAAGCTCAACAGAGAGCTGCGGCTGATGTG  
80 90 100 110 120 130 140

ACGAGGGTCAGACACCATGACGCTTCCGGGGCGGCTCGAGGTGCGCGGTGGAACCGAGACAT X  
|||||  
ACGAGGGTCAGACACCATGACGCTTCCGGGGCGGCTCGAGGTGCGCGGTGGAACCGAGACAT  
150 160 170 180 190 200

4. US-09-697-123B-5 (1-208)

US-09-697-123B-2 Sequence 22, Application US/09697123B

Initial Score = 207 Optimized Score = 207 Significance = 1.00  
Residue Identity = 99% Matches = 207 Mismatches = 1  
Gaps = 0 Conservative Substitutions = 0

X 10 20 30 40 50 60 70  
TCAAGAGAAAGCCGTACGACCTGCGCCCGCGCTGCTGCTATTAAGCTCAACAGAGAGCTGCGGCTGATGTG  
|||||  
TCAAGAGAAAGCCGTACGACCTGCGCCCGCGCTGCTGCTATTAAGCTCAACAGAGAGCTGCGGCTGATGTG  
10 20 30 40 50 60 70

GGGAGCCCATACGTCGTCGACGCTTACCCGAGAGAGAGCTGCGCTGATTAAGCTCAACAGAGAGCTGCGGCTGATGTG  
|||||  
GGGAGCCCATACGTCGTCGACGCTTACCCGAGAGAGAGCTGCGCTGATTAAGCTCAACAGAGAGCTGCGGCTGATGTG  
80 90 100 110 120 130 140

ACGAGGGTCAGACACCATGACGCTTCCGGGGCGGCTCGAGGTGCGCGGTGGAACCGAGACAT X  
|||||  
ACGAGGGTCAGACACCATGACGCTTCCGGGGCGGCTCGAGGTGCGCGGTGGAACCGAGACAT  
150 160 170 180 190 200

5. US-09-697-123B-5 (1-208)

US-09-697-123B-1 Sequence 10, Application US/09697123B

Initial Score = 186 Optimized Score = 186 Significance = 0.60  
Residue Identity = 89% Matches = 186 Mismatches = 22  
Gaps = 0 Conservative Substitutions = 0

X 10 20 30 40 50 60 70  
TCAAGAGAAAGCCGTACGACCTGCGCCCGCGCTGCTGCTATTAAGCTCAACAGAGAGCTGCGGCTGATGTG  
|||||  
TCAAGAGAAAGCCGTACGACCTGCGCCCGCGCTGCTGCTATTAAGCTCAACAGAGAGCTGCGGCTGATGTG  
10 20 30 40 50 60 70

GGGAGCCCATACGTCGTCGACGCTTACCCGAGAGAGAGCTGCGCTGATTAAGCTCAACAGAGAGCTGCGGCTGATGTG  
|||||  
GGGAGCCCATACGTCGTCGACGCTTACCCGAGAGAGAGCTGCGCTGATTAAGCTCAACAGAGAGCTGCGGCTGATGTG  
80 90 100 110 120 130 140

ACGAGGGTCAGACACCATGACGCTTCCGGGGCGGCTCGAGGTGCGCGGTGGAACCGAGACAT X  
|||||  
ACGAGGGTCAGACACCATGACGCTTCCGGGGCGGCTCGAGGTGCGCGGTGGAACCGAGACAT  
150 160 170 180 190 200

6. US-09-697-123B-5 (1-208)

US-09-697-123B-1 Sequence 11, Application US/09697123B

Initial Score = 185 Optimized Score = 185 Significance = 0.58  
Residue Identity = 88% Matches = 185 Mismatches = 23  
Gaps = 0 Conservative Substitutions = 0

X 10 20 30 40 50 60 70  
TCAAGAGAAAGCCGTACGACCTGCGCCCGCGCTGCTGCTATTAAGCTCAACAGAGAGCTGCGGCTGATGTG  
|||||  
TCAAGAGAAAGCCGTACGACCTGCGCCCGCGCTGCTGCTATTAAGCTCAACAGAGAGCTGCGGCTGATGTG  
10 20 30 40 50 60 70

GGGAGCCCATACGTCGTCGACGCTTACCCGAGAGAGAGCTGCGCTGATTAAGCTCAACAGAGAGCTGCGGCTGATGTG  
|||||  
GGGAGCCCATACGTCGTCGACGCTTACCCGAGAGAGAGCTGCGCTGATTAAGCTCAACAGAGAGCTGCGGCTGATGTG  
80 90 100 110 120 130 140

ACGAGGGTCAGACACCATGACGCTTCCGGGGCGGCTCGAGGTGCGCGGTGGAACCGAGACAT X  
|||||  
ACGAGGGTCAGACACCATGACGCTTCCGGGGCGGCTCGAGGTGCGCGGTGGAACCGAGACAT  
150 160 170 180 190 200

7. US-09-697-123B-5 (1-208)

US-09-697-123B-3 Sequence 3, Application US/09697123B

Initial Score = 185 Optimized Score = 185 Significance = 0.58  
Residue Identity = 88% Matches = 185 Mismatches = 23  
Gaps = 0 Conservative Substitutions = 0

X 10 20 30 40 50 60 70  
TCAAGAGAAAGCCGTACGACCTGCGCCCGCGCTGCTGCTATTAAGCTCAACAGAGAGCTGCGGCTGATGTG  
|||||  
TCAAGAGAAAGCCGTACGACCTGCGCCCGCGCTGCTGCTATTAAGCTCAACAGAGAGCTGCGGCTGATGTG  
10 20 30 40 50 60 70

GGGAGCCCATACGTCGTCGACGCTTACCCGAGAGAGAGCTGCGCTGATTAAGCTCAACAGAGAGCTGCGGCTGATGTG  
|||||  
GGGAGCCCATACGTCGTCGACGCTTACCCGAGAGAGAGCTGCGCTGATTAAGCTCAACAGAGAGCTGCGGCTGATGTG  
80 90 100 110 120 130 140

ACGAGGGTCAGACACCATGACGCTTCCGGGGCGGCTCGAGGTGCGCGGTGGAACCGAGACAT X  
|||||  
ACGAGGGTCAGACACCATGACGCTTCCGGGGCGGCTCGAGGTGCGCGGTGGAACCGAGACAT  
150 160 170 180 190 200

8. US-09-697-123B-5 (1-208)



Multiple sequence alignment score	=	144	Optimized Score	=	170	Significance	=	-0.21
Residue Identity	=	80%	Matches	=	180	Mismatches	=	28
Gaps	=	15	Conservative Substitutions	=	0			

GGGAGCC-----CATCAGTCGTGCAGCGCTACCGAAGAAGACGTCGTGGCCACCATCGATAATCTGTC

X  
 10  
 20  
 30  
 40  
 50  
 60  
 70  
 TTAAGGAGTAAGCGCTACGACCTGGCCCCGCGTGGTCCGCTTTAAGTCATCAAGAAGCTCGGGGTGCATATCG  
 TTAAGGAGTAAGCGCTACGATCTGGCCCCGCGTGGTCCGCTTTAAGTCATCAAGAAGCTCGGGGTGCATATCGCA







8. US-09-697-123B-4 (1-207)  
US-09-697-123B-2 Sequence 20, Application US/09697123B

35GATCCATCACCAGCTCGACGCTACCGAGAGAGAGCTGTCGCCACATGATACCTGGTCCGGCTCC  
80 90 100 110 120 130 140  
|||||  
35GAGCCCATCAGCTCGCTCGACGCTACCGAGAGAGCTGTCGCCACATGGAATATCTGTGCTCCGCTTCC  
80 90 100 110



80  
90  
100  
110  
120  
130  
140

GGGATTCGATCACCACCTCG--ACGCTACCCGAAGAGAGACTGTGCACCATCGAGTACTGTTGCGGCCG  
| | | | |  
CCCCCGCGAAGCACCTCGACACCACTAACCGAAGAGAGACTGTGCGCCACCATGAGTAGTACTGATGATCGGCG

X	10	20	30	40	50	60	70
TC	AAGAGAAGCGCTACGACCTGGCCCGGTGCCTACAAAGTCAACAAGATGGGCCCTGCATCTCG						
TCAA	GAGAGAAGCGCTACGATCTTGCCCGCGGTGCTGCTACAAAGTTGAACAAGACGTGGGCCCTGGCGGCA						
X	10	20	30	40	50	60	70

```

      80      90      100      110      120      130
GCGATCC-----GATCACCAGCTGACGCGTGCAGAGAGACGCTGCGCACCATGAGTACCTGCTCC
      |||||      |||||      |||||      |||||      |||||      |||||
CCAATCGGCTCAGGTGACCAACCACTTACCGAGAGAGACGCTGCGCACCATGAGTACCTGCTGC
      80      90      100      110      120      130      140
140      150      160      170      180      190      200
GCTTCGACGAGGCTCAGACGATGA-CGTTCCGGGGGGGACCGCGTTCGGTGGAGACCGACGACAT
|||||      |||||      |||||      |||||      |||||      |||||
GCCTGCACGAGGCGCCAGACCACTGATACCGCCCGCGCGCTCGAGGTGCGCGTGGATGTGACGACAT
150      160      170      180      190      200      210
X

```

25. US-09-697-123b-4 (1-207)  
US-09-697-123b-2 Sequence 25, Application US/09697123B

```

Initial Score      =      19      Optimized Score      =      19      Significance      =      -2.59
Residue Identity  =     100%      Matches              =      19      Mismatches      =      0
Gaps              =      0      Conservative Substitutions      =      0
X
TCAGGAGAGAGCGCTACGA
|||||
TCAGGAGAGAGCGCTACGA
X
10
X

```







GCGATCCGATCACCACTCTCACCCTTACCGAAGAACCTCTGCCACCATCGATCTGGTCCGTCC  
 GGCAGCGCATCAGTGGTCGACTCTACCGAAGAACCTCTGCCACCATCGATCTGGTCCGTCC



X            10            20            30            40            50            60            70

TCAAGAGAAGCGCTACGACTGGCCAGGCTTGCCTGAACAAGTCAACAAGCTCGGGCTCCGGCC

```
80      90      100      110      120      130
CGGATCG-----ATCACCAGCTCCAGCGTGCAGGAGAGACAGTCCGTCGACCATCGAGTACCTGCTCC
|||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
CCGAGTCGGCCGTACCCGCTGCAGCCAGCTGACCGAGAGGAGTGTGTCGCGCACCATCGAGTACCTGTGTC
80      90      100      110      120      130      140
140      150      160      170      180      190      200
GTCGTGCAGCAGGGGTGACGACAGATGATGACCGCTCCGGGGGCGACCGCAGGTTCGGTGGAGACCGAGACAT  X
|  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
GCCTGCACAGAGGCCACAGCCATGATGACGCTTCCGGCGCGCTCGAGGTGCCGTGGAGACCGAGACATAT  X
150      160      170      180      190      200      210
X
```

25. US-09-697-123B-3 (1-208)  
US-09-697-123B-2 Sequence 25, Application US/09697123B

Initial Score	=	19	Optimized Score	=	19	Significance	=	-2.69
Residue Identity	=	100%	Matches	=	19	Mismatches	=	0
Gaps	=	0	Conservative Substitutions	=	0			0

  

X	10	X
TCAGGAGAGGCGCTACGA		
TCAGGAGAGGCGCTACGA		
X	10	X

Sequence Name	Description	Init.	Opt.	length	Score	Score	sig.	Frame
---------------	-------------	-------	------	--------	-------	-------	------	-------

X	10	20	30	40	50	60	70
TC	AAGS	GAACGCT	ACAGACTGGCCCGGCTGGCCGCCTAC	AGTCAACA	CAGACGTGGTGCTTA	CAAGCTG	G
X	10	20	30	40	50	60	70
TC	AAGS	GAACGCT	ACAGACTGGCCCGGCTGGCCGCCTAC	AGTCAACA	CAGACGTGGTGCTTA	CAAGCTG	G
X	10	20	30	40	50	60	70
TC	AAGS	GAACGCT	ACAGACTGGCCCGGCTGGCCGCCTAC	AGTCAACA	CAGACGTGGTGCTTA	CAAGCTG	G

8. US-09-697-123B-2 (1-208)  
US-09-697-123B-1 Sequence 19, Application US/09697123B

GCACGCCGATACCAAGCTTCGACGCTGACCCAGGAAACCTGTACCCACATCGAGTACCTGGTGGGGGTGG  
 80 90 100 110 120 130 140  
 AGAATGTCGCGCAACCCACCGACCTGACCGAGAGAGCTGTCCCACTGCAATCAATACTGGTGGCGCTTGG  
 90 100 110 120 130 140

X 10 20 30 40 50 60 70  
 TCAAGGAGAGCGCTACGACCTGGCCCGGGTGGCCGCTACAGGTCACACAGAAGCTCGGTCTGAAACGTCG







FastDB - Fast Pairwise Comparison of Sequences  
Release 5.4

Query sequence being compared:	US-09-657-123B-1 (1-208)
Number of sequences searched:	26
Number of scores above cutoff:	24

Results of the initial comparison of US-09-697-123B-1 (1-208) with  
File : US09697123B.seq

## PARAMETERS

	Unary	K-tuple
Similarity matrix	1	4
Mismatch penalty	5.00	30
Gap penalty	0.33	30
Gap size penalty	34	208
Cutoff score	0	
Randomization group		

## SEARCH STATISTICS

Scores:	Mean	Median	Standard Deviation
---------	------	--------	--------------------

Times:	CPU	Total Elapsed
--------	-----	---------------

00:00:00:00

Number of residues: 5

Number of scores above cutoff:

The scores below are sorted by initial score. Significance is calculated based on initial score.

A 100% identical sequence to the query sequence was found:

Sequence Name	Description	Length	Init. Opt. Score	Sig. Frame
1	2	3	4	5

1. US-09-697-123B-1 Sequence 1, Application U	208	208	1.02	0
---	-----	-----	------	---

Sequence Name	Description	Length	Info.	Opt. Score	Sig.	Frame
2. US-09-697-1238-3	Sequence 3, Application U	208	200	200	0.86	0
3. US-09-697-1238-1	Sequence 15, Application	208	197	157	0.80	0
4. US-09-697-1238-1	Sequence 11, Application	208	189	169	0.64	0
5. US-09-697-1238-2	Sequence 2, Application U	208	189	159	0.64	0
6. US-09-697-1238-1	Sequence 10, Application	208	186	166	0.58	0
7. US-09-697-1238-1	Sequence 19, Application	208	184	164	0.54	0
8. US-09-697-1238-8	Sequence 8, Application U	208	183	163	0.52	0
9. US-09-697-1238-1	Sequence 17, Application	208	182	162	0.50	0
10. US-09-697-1238-1	Sequence 16, Application	208	182	162	0.50	0
11. US-09-697-1238-6	Sequence 6, Application U	208	182	162	0.50	0
12. US-09-697-1238-8	Sequence 5, Application U	208	182	162	0.50	0
13. US-09-697-1238-2	Sequence 22, Application	208	181	161	0.48	0
14. US-09-697-1238-2	Sequence 23, Application	208	181	161	0.48	0
15. US-09-697-1238-2	Sequence 24, Application	208	175	175	0.36	0
16. US-09-697-1238-2	Sequence 20, Application	205	172	184	0.30	0
17. US-09-697-1238-4	Sequence 4, Application U	207	162	189	0.10	0
18. US-09-697-1238-1	Sequence 12, Application	207	156	182	0.04	0
19. US-09-697-1238-8	Sequence 9, Application U	223	146	173	-0.22	0
20. US-09-697-1238-1	Sequence 13, Application	223	142	168	-0.30	0
21. US-09-697-1238-1	Sequence 14, Application	214	136	169	-0.42	0
22. US-09-697-1238-7	Sequence 7, Application U	214	129	168	-0.56	0
23. US-09-697-1238-1	Sequence 18, Application	211	114	162	-0.86	0
24. US-09-697-1238-2	Sequence 21, Application	214	109	170	-0.96	0

1. US-09-697-123B-1 (1-208)

US-09-69/-123B-1 Sequence 1, Application US/09697123B

Initial Score	-	208	Optimized Score	-	208	Significance	-	1.02
Residue Identity	-	100%	Matches	-	208	Mismatches	-	0
Gaps	-	0	Conservative Substitutions	-			-	0

X 10 20 30 40 50 60 70  
 TCAAGAGAGCGCTACGACCTGGCCGGGAGGCGCCTTCAAGGTCAACAAAGAACTCGGCTGCACTCG  
 TCAAGGAGAGCGCTACGACCTGGCCGGGAGGCGCCTTCAAGGTCAACAAAGAACTCGGCTGCACTCG  
 X 10 20 30 40 50 60 70

GCATCGATCACCACCTCTCCAGTCGCGGAAACGTCGTGGCCACATCAGTACTGTCCGGCTGG  
80 90 100 110 120 130 140  
GCGATCCGATCACCACCTCCACGTCGCGGAAAGCGTCGCCACATCAGTACTGTCCGGCTGG  
80 90 100 110 120 130 140

150  
 160  
 170  
 180  
 190  
 200  
 X

2. US-09-697-123B-1 (1-208)

Initial Score	-	200	Optimized Score	-	200	Significance	-	0.86
Residue Identity	-	96%	Matches	-	200	Mismatches	-	8
Gaps	-	0	Conservative Substitutions	-			-	0

[illegible]

```
|||||
GCGATCCGATCACAGCTCCACCTGACCCCAAGACGCTGCTGCGCACCATGATGATCTGCTCGCTGC
80      90      100      110      120      130      140
150      160      170      180      190      200      X
ACGAGGGCCAGACACAGATACCGCTCCCGGGCGGACCGAGGTGCGGTTGACACGACGACAT
|||||
ACGAGGGTCAAGCAGACGATGACCGTCCGGGGCGGACCGAGGTCTCCGGTGTGAGACCGGACGACAT
150      160      170      180      190      200      X
```

## 3. US-09-697-123b-1 (1-208)

US-09-697-123b-1 Sequence 15, Application US/09697123b

```
Initial Score = 197 Optimized Score = 197 Significance = 0.80
Residue Identity = 94% Matches = 197 Mismatches = 11
Gaps = 0 Conservative Substitutions = 0
X
TCAAGGAGAGCGCTACGACCTGCGGGGCTAGCGCCGCTACAGATCAAGAAAGCTGCGGCTGCAGCTGC
|||||
TCAAGGAGAGCGCTACGACCTGCGGGGCTAGCGCCGCTACAGATCAAGAAAGCTGCGGCTGCAGCTGC
10      20      30      40      50      60      70
X      10      20      30      40      50      60      70
TCAAGGAGAGCGCTACGACCTGCGGGGCTAGCGCCGCTACAGATCAAGAAAGCTGCGGCTGCAGCTGC
80      90      100      110      120      130      140
GCGATCCGATCACAGCTCCACGCTGACCGAGAGAGAGCTGCTGCGCCATGATGATGATGATGATGATG
|||||
GTGAGCGGATCACAGCTCCACGCTGACCGAGAGAGAGCTGCTGCGCCATGATGATGATGATGATGATG
80      90      100      110      120      130      140
```

```
150      160      170      180      190      200      X
ACGAGGGCCAGACACAGATGACCGCTCCCGGGCGGACCGAGGTGCGGTTGACGACGACAT
|||||
ACGAGGGTCAAGCAGACGATGACCGTCCGGGGCGGACCGAGGTCTCCGGTGTGAGACCGGACGACAT
150      160      170      180      190      200      X
```

## 4. US-09-697-123b-1 (1-208)

US-09-697-123b-1 Sequence 11, Application US/09697123b

```
Initial Score = 189 Optimized Score = 189 Significance = 0.64
Residue Identity = 90% Matches = 189 Mismatches = 19
Gaps = 0 Conservative Substitutions = 0
X
TCAAGGAGAGCGCTACGACCTGCGGGGCTAGCGCCGCTACAGATCAAGAAAGCTGCGGCTGCAGCTGC
|||||
TCAAGGAGAGCGCTACGACCTGCGGGGCTAGCGCCGCTACAGATCAAGAAAGCTGCGGCTGCAGCTGC
10      20      30      40      50      60      70
X      10      20      30      40      50      60      70
TCAAGGAGAGCGCTACGACCTGCGGGGCTAGCGCCGCTACAGATCAAGAAAGCTGCGGCTGCAGCTGC
80      90      100      110      120      130      140
GCGATCCGATCACAGCTCCACGCTGACCGAGAGAGAGCTGCTGCGCCATGATGATGATGATGATGATG
|||||
GCCAGCCCATCACAGCTCCACGCTGACCGAGAGAGAGCTGCTGCGCCATGATGATGATGATGATGATG
80      90      100      110      120      130      140
```

```
150      160      170      180      190      200      X
ACGAGGGCCAGACACAGATGACCGCTCCCGGGCGGACCGAGGTGCGGTTGACGACGACAT
|||||
ACGAGGGTCAAGCAGACGATGACCGTCCGGGGCGGACCGAGGTCTCCGGTGTGAGACCGGACGACAT
150      160      170      180      190      200      X
```

## 5. US-09-697-123b-1 (1-208)

US-09-697-123b-2 Sequence 2, Application US/09697123b

```
Initial Score = 189 Optimized Score = 189 Significance = 0.64
Residue Identity = 90% Matches = 189 Mismatches = 19
Gaps = 0 Conservative Substitutions = 0
X
TCAAGGAGAGCGCTACGACCTGCGGGGCTAGCGCCGCTACAGATCAAGAAAGCTGCGGCTGCAGCTGC
|||||
TCAAGGAGAGCGCTACGACCTGCGGGGCTAGCGCCGCTACAGATCAAGAAAGCTGCGGCTGCAGCTGC
10      20      30      40      50      60      70
X      10      20      30      40      50      60      70
TCAAGGAGAGCGCTACGACCTGCGGGGCTAGCGCCGCTACAGATCAAGAAAGCTGCGGCTGCAGCTGC
80      90      100      110      120      130      140
```

```
TCAAGGAGAGCGCTACGACCTGCGGGGCTAGCGCCGCTACAGATCAAGAAAGCTGCGGCTGCAGCTGC
X      10      20      30      40      50      60      70
```

```
80      90      100      110      120      130      140
GCGATCCGATCACAGCTCCACGCTGACCGAGAGAGAGCTGCTGCGCCATGATGATGATGATGATGATG
|||||
GCAAGCCGATCACAGCTCCACGCTGACCGAGAGAGAGCTGCTGCGCCATGATGATGATGATGATGATG
80      90      100      110      120      130      140
```

```
150      160      170      180      190      200      X
ACGAGGGCCAGACACAGATGACCGCTCCCGGGCGGACCGAGGTGCGGTTGACGACGACAT
|||||
ACGAGGGTCAAGCAGACGATGACCGTCCGGGGCGGACCGAGGTCTCCGGTGTGAGACCGGACGACAT
150      160      170      180      190      200      X
```

## 6. US-09-697-123b-1 (1-208)

US-09-697-123b-1 Sequence 10, Application US/09697123b

```
Initial Score = 186 Optimized Score = 186 Significance = 0.58
Residue Identity = 89% Matches = 186 Mismatches = 22
Gaps = 0 Conservative Substitutions = 0
X
TCAAGGAGAGCGCTACGACCTGCGGGGCTAGCGCCGCTACAGATCAAGAAAGCTGCGGCTGCAGCTGC
|||||
TCAAGGAGAGCGCTACGACCTGCGGGGCTAGCGCCGCTACAGATCAAGAAAGCTGCGGCTGCAGCTGC
10      20      30      40      50      60      70
X      10      20      30      40      50      60      70
TCAAGGAGAGCGCTACGACCTGCGGGGCTAGCGCCGCTACAGATCAAGAAAGCTGCGGCTGCAGCTGC
80      90      100      110      120      130      140
```

```
80      90      100      110      120      130      140
GCGATCCGATCACAGCTCCACGCTGACCGAGAGAGAGCTGCTGCGCCATGATGATGATGATGATGATG
|||||
GCCAGCCCATCACAGCTCCACGCTGACCGAGAGAGAGCTGCTGCGCCATGATGATGATGATGATGATG
80      90      100      110      120      130      140
```

```
150      160      170      180      190      200      X
ACGAGGGCCAGACACAGATGACCGCTCCCGGGCGGACCGAGGTGCGGTTGACGACGACAT
|||||
ACGAGGGTCAAGCAGACGATGACCGTCCGGGGCGGACCGAGGTCTCCGGTGTGAGACCGGACGACAT
150      160      170      180      190      200      X
```

## 7. US-09-697-123b-1 (1-208)

US-09-697-123b-1 Sequence 19, Application US/09697123b

```
Initial Score = 184 Optimized Score = 184 Significance = 0.54
Residue Identity = 88% Matches = 184 Mismatches = 24
Gaps = 0 Conservative Substitutions = 0
X
TCAAGGAGAGCGCTACGACCTGCGGGGCTAGCGCCGCTACAGATCAAGAAAGCTGCGGCTGCAGCTGC
|||||
TCAAGGAGAGCGCTACGACCTGCGGGGCTAGCGCCGCTACAGATCAAGAAAGCTGCGGCTGCAGCTGC
10      20      30      40      50      60      70
X      10      20      30      40      50      60      70
TCAAGGAGAGCGCTACGACCTGCGGGGCTAGCGCCGCTACAGATCAAGAAAGCTGCGGCTGCAGCTGC
80      90      100      110      120      130      140
```

```
80      90      100      110      120      130      140
GCGATCCGATCACAGCTCCACGCTGACCGAGAGAGAGCTGCTGCGCCATGATGATGATGATGATGATG
|||||
GCCAGCCCATCACAGCTCCACGCTGACCGAGAGAGAGCTGCTGCGCCATGATGATGATGATGATGATG
80      90      100      110      120      130      140
```

```
150      160      170      180      190      200      X
ACGAGGGCCAGACACAGATGACCGCTCCCGGGCGGACCGAGGTGCGGTTGACGACGACAT
|||||
ACGAGGGTCAAGCAGACGATGACCGTCCGGGGCGGACCGAGGTCTCCGGTGTGAGACCGGACGACAT
150      160      170      180      190      200      X
```

## 8. US-09-697-123b-1 (1-208)

US-09-697-123b-8 Sequence 8, Application US/09697123b

```
Initial Score = 183 Optimized Score = 183 Significance = 0.52
Residue Identity = 87% Matches = 183 Mismatches = 25
```

150 160 170 180 190 200 X  
ACGAGGCCACACACGATGACCGTCCCGGGCGGCACCGAGGTGCCGTTGAGACCGACGACAT

|||||  
ATGAGGCTCATCGACGCTACGCTTCCAGGTGCGAGGTGCCATGATACGACGACAT  
150 160 170 180 190 200 X

14. US-09-697-123B-1 (1-208)  
US-09-697-123B-2 Sequence 22, Application US/09697123B

Initial Score = 181 Optimized Score = 181 Significance = 0.48  
Residue Identity = 87% Matches = 181 Mismatches = 27  
Gaps = 0 Conservative Substitutions = 0

X  
TCAAGAGAGCGCTACGACCTGCGCCGGGTAGCCGCTTCAAGGTCAACAGAGCTGCGCTTCCAGCTG  
|||||  
TCAAGAGAGAGCGCTACGACCTGCGCCGGGTAGCCGCTTCAAGGTCAACAGAGCTGCGCTTCCAGCTG  
10 20 30 40 50 60 70  
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140  
GCGATCCGATCACCAGCTCCAGCTGACCGAGAGAGAGCTGCTGCGACCATCGAGTACTGCTGCTGCTG  
|||||  
GCGATCCGATCACCAGCTCCAGCTGACCGAGAGAGAGCTGCTGCGACCATCGAGTACTGCTGCTGCTG  
80 90 100 110 120 130 140

150 160 170 180 190 200 X  
ACGAGGCGCCAGCACACGATGATGCTGCGCGGCGCACCGAGGTGCGGCTTGAAGACCGACGACAT  
|||||  
ACGAGGCTCAGACGATGATGCTGCGCGGCGCACCGAGGTGCGGCTTGAAGACCGACGACAT  
150 160 170 180 190 200 X

15. US-09-697-123B-1 (1-208)  
US-09-697-123B-2 Sequence 24, Application US/09697123B

Initial Score = 175 Optimized Score = 175 Significance = 0.36  
Residue Identity = 84% Matches = 175 Mismatches = 33  
Gaps = 0 Conservative Substitutions = 0

X  
TCAAGAGAGCGCTACGACCTGCGCGGCGTACGACGCTTCAAGGTCAACAGAGCTGCGCTTCCAGCTG  
|||||  
TCAAGAGAGCGCTACGACCTGCGCGGCGTACGACGCTTCAAGGTCAACAGAGCTGCGCTTCCAGCTG  
10 20 30 40 50 60 70  
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140  
GCGATCCGATCACCAGCTCCAGCTGACCGAGAGAGAGCTGCTGCGACCATCGAGTACTGCTGCTGCTG  
|||||  
GCGATCCGATCACCAGCTCCAGCTGACCGAGAGAGAGCTGCTGCGACCATCGAGTACTGCTGCTGCTG  
80 90 100 110 120 130 140

150 160 170 180 190 200 X  
ACGAGGCGCCAGCACACGATGATGCTGCGCGGCGCACCGAGGTGCGGCTTGAAGACCGACGACAT  
|||||  
ACGAGGCGCCAGCACACGATGATGCTGCGCGGCGCACCGAGGTGCGGCTTGAAGACCGACGACAT  
150 160 170 180 190 200 X

16. US-09-697-123B-1 (1-208)  
US-09-697-123B-2 Sequence 20, Application US/09697123B

Initial Score = 172 Optimized Score = 184 Significance = 0.30  
Residue Identity = 91% Matches = 190 Mismatches = 15  
Gaps = 3 Conservative Substitutions = 0

X  
TCAAGAGAGCGCTACGACCTGCGCGGCGTACGACGCTTCAAGGTCAACAGAGCTGCGCTTCCAGCTG  
|||||  
TCAAGAGAGCGCTACGACCTGCGCGGCGTACGACGCTTCAAGGTCAACAGAGCTGCGCTTCCAGCTG  
10 20 30 40 50 60 70  
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140  
GCGATCCGATCACCAGCTCCAGCTGACCGAGAGAGAGCTGCTGCGACCATCGAGTACTGCTGCTGCTG  
|||||  
GCGATCCGATCACCAGCTCCAGCTGACCGAGAGAGAGCTGCTGCGACCATCGAGTACTGCTGCTGCTG  
80 90 100 110 120 130 140

GCGATCCGATCACCAGCTCCAGCTGACCGAGAGAGAGCTGCTGCGACCATCGAGTACTGCTGCTGCTG  
80 90 100 110 120 130 140

150 160 170 180 190 200 X  
ACGAGGCGCCAGCACACGATGATGCTGCGCGGCGCACCGAGGTGCGGCTTGAAGACCGACGACAT  
|||||  
ACGAGGCGCCAGCACACGATGATGCTGCGCGGCGCACCGAGGTGCGGCTTGAAGACCGACGACAT  
150 160 170 180 190 200 X

17. US-09-697-123B-1 (1-208)  
US-09-697-123B-4 Sequence 4, Application US/09697123B

Initial Score = 162 Optimized Score = 189 Significance = 0.10  
Residue Identity = 93% Matches = 194 Mismatches = 13  
Gaps = 1 Conservative Substitutions = 0

X  
TCAAGAGAGCGCTACGACCTGCGCGGCGTACGACGCTTCAAGGTCAACAGAGCTGCGCTTCCAGCTG  
|||||  
TCAAGAGAGCGCTACGACCTGCGCGGCGTACGACGCTTCAAGGTCAACAGAGCTGCGCTTCCAGCTG  
10 20 30 40 50 60 70  
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140  
GCGATCCGATCACCAGCTCCAGCTGACCGAGAGAGAGCTGCTGCGACCATCGAGTACTGCTGCTGCTG  
|||||  
GCGATCCGATCACCAGCTCCAGCTGACCGAGAGAGAGCTGCTGCGACCATCGAGTACTGCTGCTGCTG  
80 90 100 110 120 130 140

150 160 170 180 190 200 X  
ACGAGGCGCCAGCACACGATGATGCTGCGCGGCGCACCGAGGTGCGGCTTGAAGACCGACGACAT  
|||||  
ACGAGGCTCAGACGATGATGCTGCGCGGCGCACCGAGGTGCGGCTTGAAGACCGACGACAT  
150 160 170 180 190 200 X

18. US-09-697-123B-1 (1-208)  
US-09-697-123B-1 Sequence 12, Application US/09697123B

Initial Score = 159 Optimized Score = 182 Significance = 0.04  
Residue Identity = 89% Matches = 187 Mismatches = 20  
Gaps = 1 Conservative Substitutions = 0

X  
TCAAGAGAGCGCTACGACCTGCGCGGCGTACGACGCTTCAAGGTCAACAGAGCTGCGCTTCCAGCTG  
|||||  
TCAAGAGAGCGCTACGACCTGCGCGGCGTACGACGCTTCAAGGTCAACAGAGCTGCGCTTCCAGCTG  
10 20 30 40 50 60 70  
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140  
GCGATCCGATCACCAGCTCCAGCTGACCGAGAGAGAGCTGCTGCGACCATCGAGTACTGCTGCTGCTG  
|||||  
GCGATCCGATCACCAGCTCCAGCTGACCGAGAGAGAGCTGCTGCGACCATCGAGTACTGCTGCTGCTG  
80 90 100 110 120 130 140

150 160 170 180 190 200 X  
ACGAGGCGCCAGCACACGATGATGCTGCGCGGCGCACCGAGGTGCGGCTTGAAGACCGACGACAT  
|||||  
ACGAGGCGCCAGCACACGATGATGCTGCGCGGCGCACCGAGGTGCGGCTTGAAGACCGACGACAT  
150 160 170 180 190 200 X

19. US-09-697-123B-1 (1-208)  
US-09-697-123B-9 Sequence 9, Application US/09697123B

Initial Score = 146 Optimized Score = 173 Significance = -0.22  
Residue Identity = 83% Matches = 187 Mismatches = 21  
Gaps = 15 Conservative Substitutions = 0

X  
TCAAGAGAGCGCTACGACCTGCGCGGCGTACGACGCTTCAAGGTCAACAGAGCTGCGCTTCCAGCTG  
|||||  
TCAAGAGAGCGCTACGACCTGCGCGGCGTACGACGCTTCAAGGTCAACAGAGCTGCGCTTCCAGCTG  
10 20 30 40 50 60 70  
X 10 20 30 40 50 60 70

X 10 20 30 40 50 60 70  
 TCAAGAGAGCCCTCAGACCTGGCCGGGAGCCGCTACAACTCAACAAGCTCGCTGCACCTGC  
 |||||  
 TCAAGGAGAAGCGCTACGATCTGGCCGCGCTGGGTGCTACCAAGTGAACAAGTGGGCGCTGC  
 X 10 20 30 40 50 60 70  
 GCGATTC-----GATCACACAGCTTCACAGCTGACCCAGGAAACATCTGTCGCCACCATCGAGTACCTGGTCTC

```
140 150 160 170 180 190 200
CCATCGGGCTCAGGTGAGCACCAACCCCTCACCAGAGAAAGACGTCGTCGCCCATCGAGTACTGTGTC
80 90 100 110 120 130 140
GCCTGCACGAGGGCCAGCACACGATGACCGTCGCGGCGGCGACCGAGGTGCCGGTTGAGACCGAGACAT
140 150 160 170 180 190 200
GCCTGCACGAGGGCCAGCACACGATGACCGTCGCGGCGGCGACCGAGGTGCCGGTTGAGACCGAGACAT
150 160 170 180 190 200 210
GCCTGCACGAGGGCCAGCACACGATGACCGTCGCGGCGGCGACCGAGGTGCCGGTTGAGACCGAGACAT
150 160 170 180 190 200 210
X
X
X
```



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 12, 2002, 16:59:42 : Search time 106.647 Seconds  
(Without alignments)  
3037.202 Million cell updates/sec

Title: US-09-697-123b-26

Sequence: 1 ggaatgacagcagcgtcgc 20

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST: \*  
1: em\_estdb: \*  
2: em\_esthum: \*  
3: em\_estlin: \*  
4: em\_estnu: \*  
5: em\_estov: \*  
6: em\_estpl: \*  
7: em\_estro: \*  
8: em\_hic: \*  
9: gb\_est1: \*  
10: gb\_est2: \*  
11: gb\_hic: \*  
12: gb\_est3: \*  
13: gb\_est4: \*  
14: gb\_est5: \*  
15: em\_estfun: \*  
16: em\_estom: \*  
17: gb\_gss: \*  
18: em\_gss\_hum: \*  
19: em\_gss\_inv: \*  
20: em\_gss\_pin: \*  
21: em\_gss\_vrt: \*  
22: em\_gss\_fun: \*  
23: em\_gss\_mam: \*  
24: em\_gss\_mus: \*  
25: em\_gss\_other: \*  
26: em\_gss\_pro: \*  
27: em\_gss\_tod: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	19	95.0	414	17	A2606523
2	18.4	92.0	377	9	AA143959
3	18.4	92.0	437	10	AA143959
4	18.4	92.0	502	9	AA143959
5	18.4	92.0	579	9	AA143959
6	18.4	92.0	585	13	AA143959

C 7	18.4	92.0	597	14	B0554372
C 8	18.4	92.0	638	10	AA272508
C 9	18.4	92.0	706	14	BM943312
C 10	18.4	92.0	738	13	B1697529
C 11	18.4	92.0	759	12	B0385603
C 12	18.4	92.0	768	13	B1688003
C 13	18.4	92.0	778	13	B1695451
C 14	18.4	92.0	824	13	B1155204
C 15	18.4	92.0	4072	11	BC028426
C 16	18.4	92.0	314	17	A2276243
C 17	17.4	87.0	279	9	AA144678
C 18	17.4	87.0	532	17	AA0732019
C 19	17.4	85.0	306	10	AA114205
C 20	17.4	85.0	950	17	CNS0442C
C 21	16.8	84.0	271	17	CNS024LB
C 22	16.8	84.0	272	10	BB231701
C 23	16.8	84.0	380	12	BC303432
C 24	16.8	84.0	417	13	B1297125
C 25	16.8	84.0	461	10	B1866562
C 26	16.8	84.0	477	12	BF554685
C 27	16.8	84.0	530	13	B1843127
C 28	16.8	84.0	538	10	AA175549
C 29	16.8	84.0	550	13	B1671235
C 30	16.8	84.0	554	17	BH035861
C 31	16.8	84.0	641	14	BQ904110
C 32	16.8	84.0	652	10	AA915754
C 33	16.8	84.0	661	13	B1193716
C 34	16.4	82.0	302	13	B1049959
C 35	16.4	82.0	459	9	A1260450
C 36	16.4	82.0	520	17	A2796561
C 37	16.4	82.0	521	9	AA698485
C 38	16.4	82.0	538	13	BM405728
C 39	16.4	82.0	571	17	AA0668780
C 40	16.4	82.0	602	12	BF153712
C 41	16.4	82.0	627	17	AA0757049
C 42	16.4	82.0	680	12	BC590263
C 43	16.4	82.0	703	13	B1767119
C 44	16.4	82.0	732	12	BC886620
C 45	16.4	82.0	776	13	BM408140

#### ALIGNMENTS

B0554372	H4027G03-
AA272508	xu20b08.x
BM943312	UT-M-ER0P
B1697529	603349086
BC385603	602453904
B1688003	603315768
B1695451	603346455
B1155204	602903188
BC028426	Hom0_snp1
A2276243	RPC1-23-1
AA144678	mr69c07.r
AA0732019	HS_5555_A
AA114205	rs70a01.Y
AL314108	Tetradon
AL180920	Tetradon
BB231701	BB231701
BC303432	fl5f03.x
B1297125	UT-R-DK0-
B1866562	fl5f05.Y
BF554685	UT-R-CO-1
B1843127	fl5f05.x
AA175549	fl5f05.x
B1671235	fl5f05.x
BH035861	RPC1-24-3
BQ904110	Ta03_2090
AA915754	EST347058
B1193716	602946680
B1049959	CM7-CM029
A1260450	LP04321.5
A2796561	2M0052A23
AA698485	HL04775.5
BM405728	EST580055
AA0668780	HS_5435.A
BF153712	046F08.Ma
AA0757049	HS_5448.B
BC590263	EST498093
B1767119	603054180
BC886620	EST512471
BM408140	EST582467

RESULT 1  
LOCUS A2606523 414 bp DNA linear GSS 13-DEC-2000  
DEFINITION 1M0428B22F Mouse 10kb plasmid U0CC1M library Mus musculus genomic  
clone U0CC1M0428B22 F, DNA sequence.  
ACCESSION A2606523  
VERSION A2606523.1 GI:11728629  
KEYWORDS GSS.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 414)  
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamill,C.,  
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Rellly,  
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.  
and Wright,D., Weiss,R.  
Mouse whole genome scaffolding with paired end reads from 10kb  
plasmid inserts  
Unpublished (2000)  
CONTACT Contact: Robert B. Weiss  
UNIVERSITY University of Utah  
INSTITUTION University of Utah Genome Center  
ADDRESS Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT  
84112, USA  
TEL: 801 585 5606  
FAX: 801 585 7177  
EMAIL: ddunn@genetics.utah.edu  
Insert length: 10000 Std Error: 0.00

JOURNAL COMMENT

## FEATURES

Plate: 0428 row: B column: 22  
Seq primer: CCTGTAAACGACGCCACG  
Class: plasmid ends  
High quality sequence stop: 414.  
Location/Qualifiers  
1. .414

/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="UUCGCM0428B2"  
/clone\_lib="Mouse 10kb plasmid UUCGCM library"  
/sex="Male"  
/note="Vector: PWD42nv: Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource  
(http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PWD42 (g147321149b1AF129072.1), a copy-number inducible derivative of plasmid RL. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adaptored vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

## BASE COUNT

81 a 100 c 87 g 146 t

Query Match 95.0%; Score 19; DB 17; Length 414;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 GATGTTGATCAGGCTGTC 20  
|||||  
Db 383 GATGTTGATCAGGCTGTC 401

RESULT 2  
AA143959/c 377 bp mRNA linear EST 12-FEB-1997  
LOCUS m13d11.r1 StrataGene mouse skin (4937313) Mus musculus cDNA clone  
DEFINITION IMAGE:606837 5 similar to TR:G498013 G498013 X104. ; mRNA

ACCESSION AA143959.1 GI:1713346  
VERSION AA143959  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus

REFERENCE  
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 377)  
Marr, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M., Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Thelning, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R.

TITLE The WashU-HMI Mouse EST Project  
JOURNAL Unpublished (1996)  
COMMENT Contact: Maria M/Mouse EST Project  
WashU-HMI Mouse EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: mouseest@wustl.edu

This clone is available royalty-free through LNL; contact the

## FEATURES

IMAGE Consortium (info@image.llnl.gov) for further information.  
MGI:372269  
Possible reversed clone: similarity on wrong strand  
Seq primer: -28m13 rev1 ET from Amersham  
High quality sequence stop: 331.  
Location/Qualifiers  
1. .377

/organism="Mus musculus"  
/strain="C57BL/6"  
/db\_xref="taxon:10090"  
/clone="IMAGE:606837"  
/clone\_lib="Stratagene mouse skin (#937313)"  
/sex="females"  
/tissue\_type="whole skin"  
/dev\_stage="11 weeks old"  
/note="Organ: skin; Vector: pBluescript SK-; Site\_1: EcoRI  
Site\_2: XhoI; Cloned unidirectionally. Primer: Oligo  
dr. whole skin from 11 week old C57BL/6 female mice.  
Average insert size: 1.0 kb; Uni-ZAP XR Vector; -5'  
adaptor sequence: 5' GAATTCGCGACGAG 3' -3' adaptor  
sequence: 5' CTCGAGTTT TTT TTT TTT TTT TTT 3' "

## BASE COUNT

104 a 107 c 102 g 64 t

Query Match 92.0%; Score 18.4; DB 9; Length 377;  
Best Local Similarity 95.0%; Pred. No. 2e+02;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 GCATGTTGATCAGGCTGTC 20  
|||||  
Db 54 GCATGTTGATCAGGCTGTC 35

RESULT 3  
AW211046/c 437 bp mRNA linear EST 03-DEC-1999  
LOCUS uo90e08.y1 NCI\_CGAP\_Mam3 Mus musculus cDNA clone IMAGE:2649830 5'  
DEFINITION similar to TR:Q92001 Q92001 TIGHT JUNCTION PROTEIN ZO-2. ; mRNA

ACCESSION AW211046  
VERSION AW211046  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 437)  
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: c9apbs-r@mail.nih.gov  
Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.  
cDNA Library Preparation: Life Technologies, Inc.  
DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)  
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:  
www.bio.llnl.gov/dbfp/image/image.html

MGI:1030282  
Seq primer: -40RP from Gibco  
High quality sequence stop: 285.  
Location/Qualifiers  
1. .437

/organism="Mus musculus"  
/strain="129 C57BL/6J FVB/N"  
/db\_xref="taxon:10090"  
/clone="IMAGE:2649830"  
/clone\_lib="NCI\_CGAP Mam3"  
/tissue\_type="tumor, gross tissue"

```

/dev-stage="10 months"
/lab-host="DH10B"
/Note="Organ: mammary; Vector: pCMV-SPORT6; Site:1: Saliv;
Site:2: Notti; Cloned unidirectionally. Primer: Oligo dt.
Library constructed by Life Technologies. Investigators
providing samples: Lothar Hennighausen/Chu-Xia Deng, NIH
Reference for transgenic model: Xu et al., Nature Genetics
22, 37-43 (1999)."
BASE COUNT      136 a 102 c 111 g      86 t      2 others
ORIGIN

Query Match      92.0%; Score 18.4; DB 10; Length 437;
Best Local Similarity 95.0%; Pred. No. 2.1e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY      1 GGATGTCATCAGGCTCTGC 20
|||||
DB      306 GGATGTCATCAGGCTCTGC 287

RESULT 4
AL045401/c      502 bp      mRNA      linear      EST 29-FEB-2000
LOCUS      DKFZp334D055.r1 434 (synonym: htes3) Homo sapiens cDNA clone
DEFINITION      DKFZp334D055 5', mRNA sequence.
ACCESSION      AL045401
VERSION      AL045401.1 GI:5433550
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.
1 (bases 1 to 502)
Duesterhoeft,A., Lauber,J., Mewes,H.W., Gassenhuber,J. and Wiemann
,S.
EST (Duesterhoeft, et al.)
JOURNAL      Unpublished (1999)
COMMENT      Contact: Duesterhoeft A
MIPS
Am Klopferplatz 18a D-82152 Martinsried, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by Qiagen (Hilden/Germany) within the cDNA sequencing
consortium of the German Genome Project.
No sl sequence available.
This clone (DKFZp334D055) is available at the RZPD in Berlin.
Please contact the RZPD: Reesourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES
Source
1..502
location/Qualifiers
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /clone="DKFZp334D055"
    /clone_lib="434 (synonym: htes3)"
    /tissue-type="testis"
    /dev-stage="adult"
    /lab_host="DH10B"
    /Note="Vector: pSPORT1; Site:1: Notti; Site:2: Saliv"
BASE COUNT      152 a 123 c 122 g      101 t      4 others
ORIGIN

Query Match      92.0%; Score 18.4; DB 9; Length 502;
Best Local Similarity 95.0%; Pred. No. 2.2e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY      1 GGATGTCATCAGGCTCTGC 20
|||||
DB      30 GGATGTCATCAGGCTCTGC 11

RESULT 5
AL045401/c
A1603914/c

```

```

LOCUS      A1603914      579 bp      mRNA      linear      EST 15-MAR-2000
DEFINITION      ms13d11.y1 Stratagene mouse skin (#937313) Mus musculus cDNA clone
IMAGE:606837 5' similar to TR:P70625 P70625 ZONLIA OCCUDENS 2
ACCESSION      A1603914
VERSION      A1603914
KEYWORDS      A1603914.1 GI:4613076
EST.
SOURCE      house mouse.
ORGANISM      Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 579)
Marrin,M., Haller,L., Kucaba,T., Martin,D., Beck,C., Wylie,T.,
Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person
,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter
,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McEann,R.,
Waterson,R. and Wilson,R.
The Mashu-NCI Mouse EST Project 1999
Unpublished (1999)
Contact: Marrin M/Mashu-NCI Mouse EST Project 1999
Washington University School of Medicine
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@waterson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
This read is a RESEQUENCE of a previously sequenced mouse clone
This read has been verified (found to hit its original self in the
correct orientation)
Possible reversed clone: similarity on wrong strand
NCI:372269
Seq primer: -40RP from Gibco
High quality sequence stop: 406
POLYA-No.
FEATURES
Source
1..579
location/Qualifiers
    /organism="Mus musculus"
    /Strain="C57BL/6"
    /db_xref="taxon:10090"
    /clone="IMAGE:606837"
    /clone_lib="Stratagene mouse skin (#937313)"
    /sex="females"
    /tissue-type="whole skin"
    /dev-stage="11 weeks old"
    /lab_host="SOLR (kanamycin resistant)"
    /Note="Organ: skin; Vector: pBluescript SK-; Site:1: Ecoret
; Site:2: XhoI; Cloned unidirectionally. Primer: Oligo
dt. Whole skin from 11 week old C57BL/6 female mice.
Average insert size: 1.0 kb. Uni-ZAP XR Vector: -5'
adaptor sequence: 5' GAATTCGCGACGAG 3' -3' adaptor
sequence: 5' CTCGAGTCTTTTCTTTTCTTTT 3'"
BASE COUNT      160 a 146 c 170 g      101 t      2 others
ORIGIN

Query Match      92.0%; Score 18.4; DB 9; Length 579;
Best Local Similarity 95.0%; Pred. No. 2.4e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY      1 GGATGTCATCAGGCTCTGC 20
|||||
DB      55 GGATGTCATCAGGCTCTGC 36

RESULT 6
B1697763/c
LOCUS      B1697763      585 bp      mRNA      linear      EST 18-SEP-2001
DEFINITION      603346856f1 NCI_CGAP_Mam2 Mus musculus cDNA clone IMAGE:5374693 5',
IMAGE:5374693 5'
ACCESSION      B1697763
VERSION      B1697763.1 GI:15660392
KEYWORDS      EST.
SOURCE      house mouse.

```

ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1 (bases 1 to 585)  
 AUTHORS NIH-MGC <http://mhc.mcl.nih.gov/>.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cqabbs-r@mail.nih.gov](mailto:cqabbs-r@mail.nih.gov)  
 Tissue Procurement: Gilbert Smith, Ph.D.  
 cDNA Library Preparation: Life Technologies, Inc.  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
<http://image.llnl.gov>  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 Plate: LLAM11953 row: 0 column: 14  
 High quality sequence stop: 584.  
 Location/Qualifiers  
 1..585  
 /organism="Mus musculus"  
 /strain="FVB/N-3"  
 /db\_xref="taxon:10090"  
 /clone\_image="5374693"  
 /clone\_lib="NCI-CGAP\_Man2"  
 /tissue\_type="tumor, biopsy sample"  
 /dev\_stage="5 months"  
 /lab\_host="DH10B"  
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site\_1: SalI;  
 Site\_2: NotI; Cloned unidirectionally. Primer: Oligo dT.  
 Library constructed by Life Technologies. Investigator  
 providing samples: Gilbert Smith, NIH"

BASE COUNT 162 a 147 c 173 g 102 t 1 others  
 ORIGIN

Query Match 92.0%; Score 18.4; DB 13; Length 585;  
 Best Local Similarity 95.0%; Pred. No. 2.4e+02;  
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 GGATGTCATCAGGCTCTGC 20  
 ||||||||| |||||||||  
 Db 93 GGATGTCATCAGGCTCTGC 74

RESULT 7  
 B0554372 597 bp mRNA linear EST 20-JUN-2002  
 LOCUS H4027G03-5 NIA Mouse 7.4K cDNA Clone Set Mus musculus cDNA clone  
 DEFINITION H4027G03 5', mRNA sequence.  
 ACCESSION B0554372  
 VERSION B0554372.1 GI:21455260  
 KEYWORDS EST.  
 SOURCE house mouse.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1 (bases 1 to 597)  
 AUTHORS VanBuren, V., Plao, Y., Dudekula, D.B., Qian, Y., Carter, M.G., Martin  
 , P.R., Stagg, C.A., Bassey, U., Alba, K., Hamatani, T., Kargul, G.J.,  
 Luo, A.G. and Ko, M.S.H.  
 TITLE Assembly, verification, and initial annotation of NIA 7.4K mouse  
 cDNA clone set  
 JOURNAL Unpublished (2002)  
 COMMENT Other-ESTs: H4027G03-3  
 Contact: Yong Qian  
 Laboratory of Genetics  
 National Institute on Aging/National Institutes of Health  
 333 Cassell Drive, Suite 3000, Baltimore, MD 21224-6820, USA  
 Email: [cdna@lsun.grc.nia.nih.gov](mailto:cdna@lsun.grc.nia.nih.gov)  
 This clone set has been freely distributed to the community. Please  
 visit [http://lsun.grc.nia.nih.gov/cDNA/NIA\\_7\\_4K.html](http://lsun.grc.nia.nih.gov/cDNA/NIA_7_4K.html) for details.  
 Plate: H4027 row: G column: 03  
 Seq primer: -21M13 Reverse

High quality sequence stop: 597  
 POLYA-No.  
 Location/Qualifiers  
 1..597  
 /organism="Mus musculus"  
 /strain="C57BL/6"  
 /db\_xref="nlaEST:H4027G03-5"  
 /db\_xref="taxon:10090"  
 /clone\_image="H4027G03"  
 /clone\_lib="NIA Mouse 7.4K cDNA Clone Set"  
 /sex="mixed"  
 /dev\_stage="mixed"  
 /lab\_host="DH10B"  
 /note="Vector: pSPORT1; Site\_1: SalI; Site\_2: NotI; This  
 clone is among a rearranged set of 7,407 clones from more  
 than 20 cDNA libraries."

BASE COUNT 171 a 151 c 164 g 111 t  
 ORIGIN

Query Match 92.0%; Score 18.4; DB 14; Length 597;  
 Best Local Similarity 95.0%; Pred. No. 2.4e+02;  
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 GGATGTCATCAGGCTCTGC 20  
 ||||||||| |||||||||  
 Db 174 GGATGTCATCAGGCTCTGC 155

RESULT 8  
 AM272508 638 bp mRNA linear EST 03-JAN-2000  
 LOCUS XU20B08.x1 NCI-CGAP-Co14 Homo sapiens cDNA clone IMAGE:2800695 3'  
 DEFINITION Similar to TR:Q15883 Q15883 X104 ;, mRNA sequence.  
 ACCESSION AM272508  
 VERSION AM272508.1 GI:6659538  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 638)  
 AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 JOURNAL Unpublished (1997)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cqabbs-r@mail.nih.gov](mailto:cqabbs-r@mail.nih.gov)  
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.  
 Emmert-Buck, M.D., Ph.D.  
 cDNA Library Preparation: Life Technologies, Inc.  
 DNA Sequencing by: Greg Lennon, Ph.D.  
 Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
[www.bio.llnl.gov/bbrp/image/image.html](http://www.bio.llnl.gov/bbrp/image/image.html)

Possible reversed clone: similarity on wrong strand  
 Seq primer: 40UP from Gibco  
 High quality sequence stop: 441.  
 Location/Qualifiers  
 1..638  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_image="2800695"  
 /clone\_lib="NCI-CGAP-Co14"  
 /tissue\_type="moderately-differentiated adenocarcinoma"  
 /lab\_host="DH10B"  
 /note="Organ: colon; Vector: pCMV-SPORT6; Site\_1: SalI;  
 Site\_2: NotI; Cloned unidirectionally. Primer: Oligo dT.  
 Average insert size 1.7 kb. Life Technologies catalog #:  
 11531-019"

BASE COUNT 199 a 152 c 154 g 133 t  
 ORIGIN

Query Match 92.0%; Score 18.4; DB 10; Length 638;  
Best Local Similarity 95.0%; Pred. No. 2.5e+02;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGATGTTGATCAGGCTCTGC 20  
|||||  
Db 150 GGATGTTGATCAGGCTCTGC 131

RESULT 9  
BM943312/c 706 bp mRNA linear EST 14-MAR-2002  
LOCUS U1-M-EHOP-bvn-k-16-0-U1.1 NIH\_BMAP\_EHOP Mus musculus cDNA clone  
DEFINITION IMAGE:5694471 5', mRNA sequence.  
ACCESSION BM943312  
VERSION BM943312.1 GI:19426897  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 706)  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
Tissue Procurement: Dr. James Lin, University of Iowa  
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
This clone was contributed by the Brain Molecular Anatomy Project  
(BMAP)

FEATURES  
source  
Location/Qualifiers  
1..706  
/organism="Mus musculus"  
/strain="C57BL/6"  
/db\_xref="taxon:10090"  
/clone="IMAGE:5694471"  
/clone\_1lb="NIH\_BMAP\_EHOP"  
/tissue\_type="whole brain"  
/dev\_stage="embryo 18.5 dpc"  
/lab\_host="DH10B (TI phage resistant)"  
/note="Organ: brain; Vector: pYX-Asc; Site: 1: EcoR I;  
Site\_2: Not I; The library was constructed according to  
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,  
1996. Denatured mRNA was size fractionated on a 1% agarose  
gel. First strand cDNA synthesis was primed with an  
oligo-dT primer containing a Not I site. Double stranded  
cDNA was size selected according to mRNA size fraction,  
ligated with EcoR I adaptor, digested with Not I, and then  
cloned directionally into pYX-Asc vector. The library tag  
sequence located between the Not I site and the polyA tail  
is CAGCAGCAGC. This library was created for the  
University of Iowa Mouse Brain Molecular Anatomy Project  
(BMAP). Gene Discovery in the Developing Mouse Nervous  
System, supported by National Institute of Mental Health  
(NIMH), Hemin Chin, Ph.D., Program coordinator."

BASE COUNT 195 a 172 c 208 g 131 t

ORIGIN  
Query Match 92.0%; Score 18.4; DB 14; Length 706;  
Best Local Similarity 95.0%; Pred. No. 2.6e+02;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGATGTTGATCAGGCTCTGC 20  
|||||  
Db 33 GGATGTTGATCAGGCTCTGC 14

RESULT 10  
B1697529/c 738 bp mRNA linear EST 18-SEP-2001  
LOCUS 603349086F1 NCI\_CGAP\_Mam2 Mus musculus cDNA clone IMAGE:5376548 5',  
DEFINITION mRNA sequence.  
ACCESSION B1697529  
VERSION B1697529.1 GI:15660158  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 738)  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
Tissue Procurement: Gilbert Smith, Ph.D.  
cDNA Library Preparation: Life Technologies, Inc.  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone Distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLAM1958 row: 1 column: 21  
High quality sequence stop: 482.

FEATURES  
source  
Location/Qualifiers  
1..738  
/organism="Mus musculus"  
/strain="FVB/N-3"  
/db\_xref="taxon:10090"  
/clone="IMAGE:5376548"  
/clone\_1lb="NCI\_CGAP\_Mam2"  
/tissue\_type="tumor, biopsy sample"  
/dev\_stage="5 months"  
/lab\_host="DH10B"  
/note="Organ: mammary; Vector: pCMV-SPORT6; Site: 1: SalI;  
Site\_2: NotI; Cloned unidirectionally. Primer: Oligo dT.  
Library constructed by Life Technologies. Investigator  
providing samples: Gilbert Smith, NIH"

BASE COUNT 227 a 184 c 204 g 123 t

ORIGIN  
Query Match 92.0%; Score 18.4; DB 13; Length 738;  
Best Local Similarity 95.0%; Pred. No. 2.6e+02;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGATGTTGATCAGGCTCTGC 20  
|||||  
Db 346 GGATGTTGATCAGGCTCTGC 327

RESULT 11  
BG385603/c 759 bp mRNA linear EST 12-MAR-2001  
LOCUS 602453904F1 NIH\_MGC\_15 Homo sapiens cDNA clone IMAGE:4582297 5',  
DEFINITION mRNA sequence.  
ACCESSION BG385603  
VERSION BG385603.1 GI:13278469  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 759)  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)

Tissue Procurement: ATCC  
cDNA Library Preparation: Ling Hong/Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: NIH Intramural Sequencing Center  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLCM1304 row: 9 column: 02  
High quality sequence stop: 714.

## FEATURES

Location/Qualifiers  
1. 759

/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4582297"  
/clone\_lib="NIH-MGC\_15"  
/tissue\_type="adenocarcinoma cell line"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: colon; Vector: pOTB; Site\_1: XhoI; Site\_2:  
EcoRI; cDNA made by oligo-dt priming. Directionally  
cloned into EcoRI/XhoI sites using the following 5'  
adaptor: GGCACGAG(G). Size-selected >500bp for average  
insert size 1.8kb. Library constructed by Ling Hong in  
the laboratory of Gerald M. Rubin (University of  
California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies)"

## BASE COUNT

248 a 153 c 204 g 154 t

## ORIGIN

Query Match 92.0%; Score 18.4; DB 12; Length 759;  
Best Local Similarity 95.0%; Pred. No. 2.7e+02;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGATGTCATCAGGCTCTGC 20  
|||||  
Db 248 GGATGTCATCAGGCTCTGC 229

RESULT 12  
BI688003/c 768 bp mRNA linear EST 18-SEP-2001  
LOCUS 603315768F1 NCI\_CGAP\_Mam6 Mus musculus cDNA clone IMAGE:5355689 5',  
DEFINITION mRNA sequence.

ACCESSION BI688003  
VERSION BI688003.1 GI:15650631  
KEYWORDS EST.  
SOURCE house mouse,  
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 768)  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
Tissue Procurement: Jeffrey Green M.D.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLAM11904 row: 9 column: 18  
High quality sequence stop: 766.

## FEATURES

Location/Qualifiers  
1. 768

/organism="Mus musculus"  
/strain="FVB/N"  
/db\_xref="taxon:10090"  
/clone="IMAGE:5353689"  
/clone\_lib="NCI\_CGAP\_Mam6"  
/sex="female, virgin"  
/tissue\_type="infiltrating ductal carcinoma"

/dev\_stage="5 months"  
/lab\_host="DH10B"  
/note="Organ: mammary; Vector: pCMV-SPORT6; Site\_1: SalI;  
Site\_2: NotI; Cloned unidirectionally. Primer: Oligo dT.  
Library constructed by Life Technologies, Investigator  
providing samples: Jeffrey Green, M.D., NIH"

BASE COUNT 232 a 206 c 204 g 126 t

## ORIGIN

Query Match 92.0%; Score 18.4; DB 13; Length 768;  
Best Local Similarity 95.0%; Pred. No. 2.7e+02;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGATGTCATCAGGCTCTGC 20  
|||||  
Db 391 GGATGTCATCAGGCTCTGC 372

RESULT 13  
BI695451/c 778 bp mRNA linear EST 18-SEP-2001  
LOCUS 603346455F1 NCI\_CGAP\_Mam2 Mus musculus cDNA clone IMAGE:5374261 5',  
DEFINITION mRNA sequence.

ACCESSION BI695451  
VERSION BI695451.1 GI:15658080  
KEYWORDS EST.  
SOURCE house mouse,  
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 778)  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
Tissue Procurement: Gilbert Smith, Ph.D.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLAM11952 row: m column: 14  
High quality sequence stop: 773.

## FEATURES

Location/Qualifiers  
1. 778

/organism="Mus musculus"  
/strain="FVB/N-3"  
/db\_xref="taxon:10090"  
/clone="IMAGE:5374261"  
/clone\_lib="NCI\_CGAP\_Mam2"  
/tissue\_type="tumor, biopsy sample"  
/dev\_stage="5 months"  
/lab\_host="DH10B"  
/note="Organ: mammary; Vector: pCMV-SPORT6; Site\_1: SalI;  
Site\_2: NotI; Cloned unidirectionally. Primer: Oligo dT.  
Library constructed by Life Technologies, Investigator  
providing samples: Gilbert Smith, NIH"

BASE COUNT 232 a 208 c 208 g 130 t

## ORIGIN

Query Match 92.0%; Score 18.4; DB 13; Length 778;  
Best Local Similarity 95.0%; Pred. No. 2.7e+02;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGATGTCATCAGGCTCTGC 20  
|||||  
Db 412 GGATGTCATCAGGCTCTGC 393

RESULT 14  
BI155204/c

LOCUS B1155204 824 bp mRNA linear EST 05-JUL-2001  
 DEFINITION 602903188F1 NCI\_CGAP\_Mam3 Mus musculus cDNA clone IMAGE:5032965 5',  
 RNA sequence.  
 ACCESSION B1155204  
 VERSION B1155204.1 GI:14615205  
 KEYWORDS EST.  
 SOURCE house mouse.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 824)  
 NIH-MGC <http://mgc.ncl.nih.gov/>.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 CONTACT: Robert Strausberg, Ph.D.  
 Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
 Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.  
 cDNA Library Preparation: Life Technologies, Inc.  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: LLM11091 row: p column: 22  
 High quality sequence stop: 807.  
 Location/Qualifiers  
 1..824  
 /organism="Mus musculus"  
 /strain="129,C57BL/6J,FVB/N"  
 /db\_xref="taxon:10090"  
 /clone\_image="5032965"  
 /clone\_id="NCI\_CGAP\_Mam3"  
 /tissue\_type="tumor, gross tissue"  
 /dev\_stage="10 months"  
 /lab\_host="DH10B"  
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site: 1; Salt;  
 Site: 2; NotI; Cloned unidirectionally. Primer: Oligo dT.  
 Library constructed by Life Technologies. Investigators  
 providing samples: Lothar Hennighausen/Chu-Xia Deng, NIH  
 Reference for transgenic model: Xu et al., Nature Genetics  
 22, 37-43 (1999)."  
 22, 37-43 (1999)."  
 BASE COUNT 238 a 201 c 234 g 149 t 2 others  
 ORIGIN  
 Query Match 92.0%; Score 18.4; DB 13; Length 824;  
 Best Local Similarity 95.0%; Pred. No. 2.8e+02;  
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Oy 1 GGATGTTGATGAGGCTCTGC 20  
 ||||||||| |||||||||  
 Db 207 GGATGTTGATGAGGCTCTGC 188  
 ||||||||| |||||||||  
 RESULT 15  
 BC028426/c 4072 bp mRNA linear HTC 25-APR-2002  
 LOCUS BC028426  
 DEFINITION Homo sapiens, clone IMAGE:4820998, mRNA.  
 ACCESSION BC028426  
 VERSION BC028426.1 GI:20306365  
 KEYWORDS HTC.  
 SOURCE Homo sapiens.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 1 (bases 1 to 4072)  
 Strausberg, R.  
 TITLE Direct Submission  
 JOURNAL Submitted (23-APR-2002) National Institutes of Health, Mammalian  
 Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
 USA  
 NIH-MGC Project URL: <http://mgc.ncl.nih.gov>  
 CONTACT: MGC help desk

Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
 Tissue Procurement: Miklos Palikovits, M.D., Ph.D.  
 cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shitaki  
 Toshiyuki and Piero Carninci (RIKEN)  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Institute for Systems Biology  
<http://www.systemsbio.org>  
 contact: [amadansystemsbiology.org](mailto:amadansystemsbiology.org)  
 Anup Madan, Jessica Fahey, Erin Helton, Mark Kerteman, Anuradha  
 Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whitting  
 Clone distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Series: IRAP Plate: 46 Row: h Column: 8  
 This clone has the following problem: frame shifted.  
 Location/Qualifiers  
 1..4072  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_image="4820998"  
 /tissue\_type="Brain, hippocampus"  
 /clone\_id="NIH MGC\_95"  
 /lab\_host="DH10B"  
 /note="Vector: pBluescript"  
 BASE COUNT 1119 a 957 c 1133 g 863 t  
 ORIGIN  
 Query Match 92.0%; Score 18.4; DB 11; Length 4072;  
 Best Local Similarity 95.0%; Pred. No. 5.5e+02;  
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Oy 1 GGATGTTGATGAGGCTCTGC 20  
 ||||||||| |||||||||  
 Db 1273 GGATGTTGATGAGGCTCTGC 1254  
 ||||||||| |||||||||  
 Search completed: November 13, 2002, 04:01:15  
 Job time : 120.147 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 12, 2002, 16:05:40 : Search time 2.03468 Seconds  
(without alignments)  
3487.380 Million cell updates/sec

Title: US-09-697-123b-26

Perfect score: 20

Sequence: 1 ggaatgctgacagagcttcgc 20

Scoring table: IDENTITY\_NUC

Gapop 10.0, Gapext 1.0

Searched: 320260 seqs, 177392727 residues

Total number of hits satisfying chosen parameters: 640520

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:\*

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*
- 5: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*
- 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*
- 12: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*
- 13: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*
- 14: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	18.4	92.0	477	10	US-09-563-817-166
2	18.4	92.0	348	10	US-09-917-800A-1477
3	18.4	92.0	4530	12	US-10-044-090-357
4	16.4	82.0	291	10	US-09-815-242-7923
5	15.8	79.0	76	10	US-09-864-761-21019
6	15.8	79.0	237	10	US-09-864-761-22613
7	15.8	79.0	280	10	US-09-960-352-6294
8	15.8	79.0	334	10	US-09-864-761-4266
9	15.8	79.0	434	10	US-09-960-352-4809
10	15.8	79.0	483	10	US-09-864-761-10561
11	15.8	79.0	486	10	US-09-864-761-5853
12	15.8	79.0	508	10	US-09-864-761-27182
13	15.8	79.0	1976	10	US-09-864-761-9455
14	15.4	77.0	1992	10	US-09-815-242-7772
15	15.2	76.0	1995	10	US-09-841-132-381
16	15.2	76.0	2374	10	US-09-892-985-5
17	15.2	76.0	2596	10	US-09-823-849A-408
18	15.2	76.0	2607	10	US-09-815-242-4067
19	15.2	76.0	2880	12	US-10-044-090-361

C 20	15	75.0	1534	10	US-09-350-874-7	Sequence 7, App11
C 21	15	75.0	1856	10	US-09-350-874-3	Sequence 3, App11
C 22	14.8	74.0	157	10	US-09-864-761-24969	Sequence 24969, A
C 23	14.8	74.0	296	10	US-09-294-093B-2249	Sequence 2249, Ap
C 24	14.8	74.0	383	10	US-09-960-352-11847	Sequence 11847, A
C 25	14.8	74.0	475	10	US-09-864-761-3510	Sequence 3510, Ap
C 26	14.8	74.0	551	10	US-09-604-287A-248	Sequence 248, App
C 27	14.8	74.0	551	10	US-09-339-338-248	Sequence 248, App
C 28	14.8	74.0	551	12	US-10-007-805-248	Sequence 248, App
C 29	14.8	74.0	558	10	US-09-864-761-9973	Sequence 9973, Ap
C 30	14.8	74.0	569	10	US-09-864-761-8447	Sequence 8447, Ap
C 31	14.8	74.0	583	10	US-09-864-761-8224	Sequence 8224, Ap
C 32	14.8	74.0	613	10	US-09-998-598-1559	Sequence 1559, Ap
C 33	14.8	74.0	663	10	US-09-833-381-553	Sequence 553, App
C 34	14.8	74.0	934	10	US-09-728-952-61	Sequence 61, App1
C 35	14.8	74.0	1125	9	US-09-938-842A-2333	Sequence 2333, Ap
C 36	14.8	74.0	1167	10	US-09-940-921B-6	Sequence 6, App11
C 37	14.8	74.0	1196	10	US-09-822-830A-445	Sequence 445, App11
C 38	14.8	74.0	1197	10	US-09-840-921B-8	Sequence 8, App11
C 39	14.8	74.0	1200	10	US-09-815-242-4132	Sequence 4132, Ap
C 40	14.8	74.0	1744	10	US-09-940-921B-10	Sequence 10, App1
C 41	14.8	74.0	1778	10	US-09-962-832-128	Sequence 128, App
C 42	14.8	74.0	2383	10	US-09-925-302-317	Sequence 317, App
C 43	14.8	74.0	5156	10	US-09-735-705-130	Sequence 130, App
C 44	14.8	74.0	5156	10	US-09-850-716A-130	Sequence 130, App
C 45	14.8	74.0	5156	10	US-09-897-778-130	Sequence 130, App

#### ALIGNMENTS

RESULT 1  
US-09-563-817-166/C  
Sequence 166, Application US/09563817  
Patent No. US20020095031A1  
GENERAL INFORMATION:  
APPLICANT: Nehls, Michael C.  
APPLICANT: Zambrowicz, Brian  
TITLE OF INVENTION: Sands, Arthur T.  
TITLE OF INVENTION: Polypeptides Encoded Thereby  
FILE REFERENCE: LEX-0021-USA  
CURRENT APPLICATION NUMBER: US/09/563,817  
PRIORITY FILING DATE: 2001-06-11  
PRIOR APPLICATION NUMBER: US 60/132,343  
PRIORITY FILING DATE: 1999-05-04  
NUMBER OF SEQ ID NOS: 1008  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 166  
LENGTH: 477  
TYPE: DNA  
ORGANISM: homo sapiens  
FEATURE:  
NAME/KEY: misc-feature  
LOCATION: (1)...(477)  
OTHER INFORMATION: n = A,T,C or G  
US-09-563-817-166

Query Match 92.0%; Score 18.4; DB 10; Length 477;  
Best Local Similarity 95.0%; Pred. No. 3.7;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGATGTTGATCAGGCTCTGC 20  
II |||||||  
Db 214 GGGGTGTGATCAGGCTCTGC 195

RESULT 2  
US-09-917-800A-1477/C  
Sequence 1477, Application US/09917800A  
Patent No. US20020119462A1  
GENERAL INFORMATION:  
APPLICANT: Mendick, Donna



```

: APPLICANT: Porter, Mark
: APPLICANT: Johnson, Kory
: APPLICANT: Castle, Arthur
: APPLICANT: Elashoff, Michael
: APPLICANT: Gene Logic, Inc.
: TITLE OF INVENTION: Molecular Toxicology Modeling
: FILE REFERENCE: 44921-5038-US
: CURRENT APPLICATION NUMBER: US/09/917,800A
: CURRENT FILING DATE: 2001-07-31
: PRIOR APPLICATION NUMBER: US 60/222,040
: PRIOR FILING DATE: 2000-07-31
: PRIOR APPLICATION NUMBER: US 60/222,880
: PRIOR FILING DATE: 2000-11-02
: PRIOR APPLICATION NUMBER: US 60/290,029
: PRIOR FILING DATE: 2001-05-11
: PRIOR APPLICATION NUMBER: US 60/290,645
: PRIOR FILING DATE: 2001-05-15
: PRIOR APPLICATION NUMBER: US 60/292,336
: PRIOR FILING DATE: 2001-05-22
: PRIOR APPLICATION NUMBER: US 60/295,798
: PRIOR FILING DATE: 2001-06-06
: PRIOR APPLICATION NUMBER: US 60/297,457
: PRIOR FILING DATE: 2001-06-13
: PRIOR APPLICATION NUMBER: US 60/298,884
: PRIOR FILING DATE: 2001-06-19
: PRIOR APPLICATION NUMBER: US 60/303,459
: PRIOR FILING DATE: 2001-07-09
: NUMBER OF SEQ ID NOS: 1740
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 1477
: LENGTH: 3348
: TYPE: DNA
: ORGANISM: Rattus norvegicus
: FEATURE:
: OTHER INFORMATION: Genbank Accession No. US20020119462A1 U75916
: NAME/KEY: misc.feature
: LOCATION: (1)..(3339)
: OTHER INFORMATION: n = a or c or g or t
: US-09-917-800A-1477

Query Match          92.0%: Score 18.4: DB 10: Length 3348:
Best Local Similarity 95.0%: Pred. No. 4.6:
Matches 19: Conservative 0: Mismatches 1: Indels 0: Gaps 0:

Qy      1 GCATGTTGATCAGGCTCTGC 20
        |||
Db      60 GCATGTTGATGAGGCTCTGC 41

RESULT 3
US-10-044-090-357/c
: Sequence 357, Application US/10044090
: Patent No. US20020137081A1
: GENERAL INFORMATION:
: APPLICANT: Olga Bandman
: TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION
: FILE REFERENCE: PA-0028 US
: CURRENT APPLICATION NUMBER: US/10/044,090
: CURRENT FILING DATE: 2002-01-09
: NUMBER OF SEQ ID NOS: 850
: SOFTWARE: PERL Program
: SEQ ID NO 357
: LENGTH: 4530
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc.feature
: OTHER INFORMATION: Incyte ID No. US20020137081A1 364726.12
US-10-044-090-357

Query Match          92.0%: Score 18.4: DB 12: Length 4530:
Best Local Similarity 95.0%: Pred. No. 4.7:
Matches 19: Conservative 0: Mismatches 1: Indels 0: Gaps 0:
```

```

Qy      1 GCATGTTGATCAGGCTCTGC 20
        |||
Db      1306 GCATGTTGATGAGGCTCTGC 1287

RESULT 4
US-09-815-242-7923/c
: Sequence 7923, Application US/09815242
: Patent No. US20020061569A1
: GENERAL INFORMATION:
: APPLICANT: Haselbeck, Robert
: APPLICANT: Ohlsen, Kari L.
: APPLICANT: Zyskind, Judith W.
: APPLICANT: Wall, Daniel
: APPLICANT: Trawick, John D.
: APPLICANT: Carr, Grant J.
: APPLICANT: Yamamoto, Robert T.
: APPLICANT: Xu, H. Howard
: TITLE OF INVENTION: Identification of Essential Genes In
: FILE REFERENCE: ELITRA.011A
: CURRENT APPLICATION NUMBER: US/09/815,242
: PRIOR FILING DATE: 2001-03-21
: PRIOR APPLICATION NUMBER: 60/191,078
: PRIOR FILING DATE: 2000-03-21
: PRIOR APPLICATION NUMBER: 60/206,848
: PRIOR FILING DATE: 2000-05-23
: PRIOR APPLICATION NUMBER: 60/207,727
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: 60/242,578
: PRIOR FILING DATE: 2000-10-23
: PRIOR APPLICATION NUMBER: 60/253,625
: PRIOR FILING DATE: 2000-11-27
: PRIOR APPLICATION NUMBER: 60/257,931
: PRIOR FILING DATE: 2000-12-22
: PRIOR APPLICATION NUMBER: 60/269,308
: PRIOR FILING DATE: 2001-02-16
: NUMBER OF SEQ ID NOS: 14110
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 7923
: LENGTH: 291
: TYPE: DNA
: ORGANISM: Pseudomonas aeruginosa
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (1)..(291)
US-09-815-242-7923

Query Match          82.0%: Score 16.4: DB 10: Length 291:
Best Local Similarity 94.4%: Pred. No. 32:
Matches 17: Conservative 0: Mismatches 1: Indels 0: Gaps 0:

Qy      1 GCATGTTGATCAGGCTCT 18
        |||
Db      103 GCATGTTGTTCAGGCTCT 86

RESULT 5
US-09-864-761-21019
: Sequence 21019, Application US/09864761
: Patent No. US20020048763A1
: GENERAL INFORMATION:
: APPLICANT: Penn, Sharon G.
: APPLICANT: Rank, David H.
: APPLICANT: Hanzel, David K.
: APPLICANT: Chen, Wensheng
: TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
: FILE REFERENCE: Aeonica-X-1
: CURRENT APPLICATION NUMBER: US/09/864,761
: CURRENT FILING DATE: 2001-05-23
: PRIOR APPLICATION NUMBER: US 60/180,312
```

```

: PRIOR FILING DATE: 2000-02-04
: PRIOR APPLICATION NUMBER: US 60/207,456
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: US 09/632,366
: PRIOR FILING DATE: 2000-08-03
: PRIOR APPLICATION NUMBER: GB 24263,6
: PRIOR FILING DATE: 2000-10-04
: PRIOR APPLICATION NUMBER: US 60/236,359
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: PCT/US01/00666
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00667
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00664
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00669
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00665
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00668
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00663
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00662
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00661
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00670
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: US 60/234,687
: PRIOR FILING DATE: 2000-09-21
: PRIOR APPLICATION NUMBER: US 09/608,408
: PRIOR FILING DATE: 2000-06-30
: PRIOR APPLICATION NUMBER: US 09/774,203
: PRIOR FILING DATE: 2001-01-29
: NUMBER OF SEQ ID NOS: 49117
: SOFTWARE: Annonmax Sequence Listing Engine vers. 1.1
: SEQ ID NO 21019
: LENGTH: 76
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: MAP TO AC012299.1
: OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 3
: OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.1
: OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.9
: OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.5
: OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.2
: OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.7
: OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.9
: OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.2
: OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.7
: OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 3.3
: OTHER INFORMATION: NT HIT: AF271735.1, EVALUE 4.00e-11
: OTHER INFORMATION: EST_HUMAN HIT: BE968839.1, EVALUE 5.00e-21
US-09-864-761-22613

```

```

Query Match          79.0%: Score 15.8; DB 10; Length 76;
Best Local Similarity 89.5%: Pred. No. 55;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 1 GCATGTTGATCAGGCTCTG 19
    |||||  |||||  |||||
DB 46 GCATGTTGCTCAGCTCTG 64

```

```

RESULT 6
US-09-864-761-22613
: Sequence 22613, Application US/09864761
: Patent No. US20020048763A1
: GENERAL INFORMATION:
: APPLICANT: Penn, Sharon G.
: APPLICANT: Rank, David R.

```

```

: APPLICANT: Hanzel, David K.
: APPLICANT: Chen, Wensheng
: TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
: FILE REFERENCE: Aecm1ca-X-1
: CURRENT APPLICATION NUMBER: US/09/864,761
: PRIOR FILING DATE: 2001-05-23
: PRIOR APPLICATION NUMBER: US 60/180,312
: PRIOR FILING DATE: 2000-02-04
: PRIOR APPLICATION NUMBER: US 60/207,456
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: US 09/632,366
: PRIOR FILING DATE: 2000-08-03
: PRIOR APPLICATION NUMBER: GB 24263,6
: PRIOR FILING DATE: 2000-10-04
: PRIOR APPLICATION NUMBER: PCT/US01/00665
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: US 60/236,359
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00666
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00667
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00664
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00669
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00665
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00668
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00662
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00661
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00670
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: US 60/234,687
: PRIOR FILING DATE: 2000-09-21
: PRIOR APPLICATION NUMBER: US 09/608,408
: PRIOR FILING DATE: 2000-06-30
: PRIOR APPLICATION NUMBER: US 09/774,203
: PRIOR FILING DATE: 2001-01-29
: NUMBER OF SEQ ID NOS: 49117
: SOFTWARE: Annonmax Sequence Listing Engine vers. 1.1
: SEQ ID NO 22613
: LENGTH: 237
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: MAP TO AC006329.2
: OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 9.5
: OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 6.1
: OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 7.3
: OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2
: OTHER INFORMATION: SWISSPROT HIT: Q64726, EVALUE 6.00e-27
: OTHER INFORMATION: EST_HUMAN HIT: BF677910.1, EVALUE 1.00e-116
: OTHER INFORMATION: NT HIT: X70170.1, EVALUE 1.00e-116
US-09-864-761-22613

```

```

Query Match          79.0%: Score 15.8; DB 10; Length 237;
Best Local Similarity 89.5%: Pred. No. 62;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 2 GATGTTGATCAGGCTCTGC 20
    |||||  |||||  |||||
DB 2 GATGTTGCTCAGGCTCTCC 20

```

```

RESULT 7
US-09-960-352-6294/C
: Sequence 6294, Application US/09960352
: Patent No. US20020137139A1

```

```
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; NUMBER OF SEQ ID NOS: 2001-09-24
; SEQ ID NO 6294
; LENGTH: 280
; TYPE: DNA
; ORGANISM: Bos taurus
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (115)
; OTHER INFORMATION: unsure at all n locations
; OTHER INFORMATION: Clone ID: 27-LIB3058-037-Q1-K1-G11
US-09-960-352-6294

Query Match          79.0%; Score 15.8; DB 10; Length 280;
Best Local Similarity 89.5%; Pred. No. 63;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GATGTTGATCAGGCTCTGC 20
Db 70 GATGTTGATCGGCTCTTC 52

RESULT 8
US-09-864-761-4266
; Sequence 4266, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aesomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; GENERAL INFORMATION:
```

```
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 4266
; LENGTH: 398
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC012299.1
; OTHER INFORMATION: EXPRESSED IN BR474, SIGNAL = 3
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.9
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.5
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.2
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.7
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.9
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.2
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.7
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 3.3
US-09-864-761-4266

Query Match          79.0%; Score 15.8; DB 10; Length 398;
Best Local Similarity 89.5%; Pred. No. 65;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GGATGTTGATCAGGCTCTG 19
Db 363 GGATGTTGATCAGGCTCTG 381

RESULT 9
US-09-960-352-4809/C
; Sequence 4809, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 4809
; LENGTH: 434
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 21-LIB3057-006-Q1-K1-F9
US-09-960-352-4809

Query Match          79.0%; Score 15.8; DB 10; Length 434;
Best Local Similarity 89.5%; Pred. No. 66;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GATGTTGATCAGGCTCTGC 20
Db 69 GATGTTGATCTGGGCTCTTC 51

RESULT 10
US-09-864-761-10561
; Sequence 10561, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
```

```

; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeonica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 10561
; LENGTH: 483
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL121758.4
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.91
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.91
US-09-864-761-10561

Query Match          79.0%; Score 15.8; DB 10; Length 483;
Best Local Similarity 89.5%; Pred. No. 67;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1  GCGTGTGATCAGGCTCTG 19
Db      370  GGTGTGTGATCAGGCTCG 388

RESULT 11
US-09-864-761-5853
; Sequence 5853, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
```

```

; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; FILE REFERENCE: Aeonica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 5853
; LENGTH: 486
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC006329.2
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 9.5
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 6.1
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 73
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2
US-09-864-761-5853

Query Match          79.0%; Score 15.8; DB 10; Length 486;
Best Local Similarity 89.5%; Pred. No. 67;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      2  GATGTGTGATCAGGCTCTGC 20
Db      261  GATGTGTGATCAGGCTCTCC 279

RESULT 12
US-09-864-761-27182
; Sequence 27182, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
```

```
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecomica-X-1
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 09/864,761
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/226,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 27182
; LENGTH: 508
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL121758.4
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.91
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.91
; OTHER INFORMATION: EST_HUMAN HIT: AU144395.1, EVALUATE 1.20e-01
; OTHER INFORMATION: NT HIT: AJ272265.1, EVALUATE 5.00e+00
; OTHER INFORMATION: SWISSPROT HIT: P21997, EVALUATE 1.10e+00
US-09-864-761-27182

Query Match          79.0%; Score 15.8; DB 10; Length 508;
Best Local Similarity 89.5%; Pred. No. 67;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
OY 1 GCATGTTGATCAGGCTG 19
DB 400 GGTGTTGATCAGGACTG 418
```

```
RESULT 13
US-09-864-761-3455/c
; Sequence 3455, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
```

```
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; FILE REFERENCE: Aecomica-X-1
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 09/864,761
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 3455
; LENGTH: 1976
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC006116.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.94
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.95
US-09-864-761-3455

Query Match          79.0%; Score 15.8; DB 10; Length 1976;
Best Local Similarity 89.5%; Pred. No. 77;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
OY 2 GATGTTGATCAGGCTG 20
DB 1153 GATGTTGATCAGGTAATGC 1135
```

```
RESULT 14
US-09-815-242-7772/c
; Sequence 7772, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
```

DB 393 GGATGTCATCAGCCTCTGC 412  
Search completed: November 12, 2002, 16:59:33  
Job time : 5.03468 secs

APPLICANT: Ohlsen, Karl L.  
APPLICANT: Zyskind, Judith W.  
APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John D.  
APPLICANT: Carr, Grant J.  
APPLICANT: Yamamoto, Robert T.  
APPLICANT: Xu, H. Howard  
TITLE OF INVENTION: Identification of Essential Genes In  
FILE REFERENCE: ELITRA.011A  
CURRENT APPLICATION NUMBER: US/09/815,242  
CURRENT FILING DATE: 2001-03-21  
PRIOR APPLICATION NUMBER: 60/191,078  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: 60/206,848  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/207,727  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 60/242,578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253,625  
PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: 60/257,931  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/269,308  
PRIOR FILING DATE: 2001-02-16  
NUMBER OF SEQ ID NOS: 14110  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 7772  
LENGTH: 1992  
TYPE: DNA  
ORGANISM: Pseudomonas aeruginosa  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)...(1992)  
US-09-815-242-7772

Query Match 77.0%: Score 15.4; DB 10; Length 1992;  
Best Local Similarity 94.1%: Pred. No. 1.2e+02;  
Matches 16: Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
OY 1 GGATGTCATCAGGCTC 17  
||||| |||||||  
DB 181 GGATGTCATCAGGCTC 165

RESULT 15  
US-09-841-132-381  
Sequence 381, Application US/09841132  
Patent No. US20020061848A1  
GENERAL INFORMATION:  
APPLICANT: Bhalla, Ajay  
APPLICANT: Skeiky, Yasir A.W.  
APPLICANT: Probst, Peter  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND  
FILE REFERENCE: 210121.469C8  
CURRENT APPLICATION NUMBER: US/09/841,132  
CURRENT FILING DATE: 2001-04-23  
NUMBER OF SEQ ID NOS: 599  
SOFTWARE: FastSeq for Windows Version 3.0/4.0  
SEQ ID NO 381  
LENGTH: 1995  
TYPE: DNA  
ORGANISM: Chlamydia pneumoniae  
US-09-841-132-381

Query Match 76.0%: Score 15.2; DB 10; Length 1995;  
Best Local Similarity 85.0%: Pred. No. 1.5e+02;  
Matches 17: Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
OY 1 GGATGTCATCAGGCTCTGC 20  
||||| ||||||| |||||

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 12, 2002, 15:56:36 ; Search time 2.69208 Seconds  
(without alignments)  
2286.857 Million cell updates/sec

Title: US-09-697-123B-26

Perfect score: 20

Sequence: 1 ggatgttcacgggtctgc 20

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents\_NA:\*
- 1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq:\*
  - 2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq:\*
  - 3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq:\*
  - 4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq:\*
  - 5: /cgn2\_6/ptodata/1/ina/PCTUS\_COMB.seq:\*
  - 6: /cgn2\_6/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query No.	Score	Match	Length	ID	Description
C 1	20	100.0	228	4	US-08-750-088A-69	Sequence 69, Appl
C 2	20	100.0	254	4	US-08-750-088A-37	Sequence 37, Appl
C 3	20	100.0	306	4	US-09-147-935A-1	Sequence 1, Appl
C 4	20	100.0	306	4	US-09-147-935A-3	Sequence 3, Appl
C 5	20	100.0	306	4	US-09-147-935A-4	Sequence 4, Appl
C 6	20	100.0	306	4	US-09-147-935A-5	Sequence 5, Appl
C 7	20	100.0	306	4	US-09-147-935A-8	Sequence 8, Appl
C 8	20	100.0	306	4	US-09-147-935A-9	Sequence 9, Appl
C 9	20	100.0	306	4	US-09-147-935A-10	Sequence 10, Appl
C 10	20	100.0	306	4	US-09-147-935A-11	Sequence 11, Appl
C 11	20	100.0	306	4	US-09-147-935A-12	Sequence 12, Appl
C 12	20	100.0	306	4	US-09-147-935A-13	Sequence 13, Appl
C 13	20	100.0	306	4	US-09-147-935A-14	Sequence 14, Appl
C 14	20	100.0	306	4	US-09-147-935A-15	Sequence 15, Appl
C 15	20	100.0	306	4	US-09-147-935A-17	Sequence 17, Appl
C 16	20	100.0	306	4	US-09-147-935A-18	Sequence 18, Appl
C 17	20	100.0	306	4	US-09-147-935A-22	Sequence 22, Appl
C 18	20	100.0	306	4	US-09-147-935A-27	Sequence 27, Appl
C 19	20	100.0	306	4	US-09-147-935A-28	Sequence 28, Appl
C 20	20	100.0	306	4	US-09-147-935A-29	Sequence 29, Appl
C 21	20	100.0	306	4	US-09-147-935A-30	Sequence 30, Appl
C 22	20	100.0	306	4	US-09-147-935A-31	Sequence 31, Appl
C 23	20	100.0	306	4	US-09-147-935A-32	Sequence 32, Appl
C 24	20	100.0	306	4	US-09-147-935A-33	Sequence 33, Appl
C 25	20	100.0	306	4	US-09-147-935A-35	Sequence 35, Appl
C 26	20	100.0	306	4	US-09-147-935A-36	Sequence 36, Appl
C 27	20	100.0	306	4	US-09-147-935A-37	Sequence 37, Appl

C 28	20	100.0	306	4	US-09-147-935A-38	Sequence 38, Appl
C 29	20	100.0	306	4	US-09-147-935A-39	Sequence 39, Appl
C 30	20	100.0	306	4	US-09-147-935A-40	Sequence 40, Appl
C 31	20	100.0	306	4	US-09-147-935A-43	Sequence 43, Appl
C 32	20	100.0	306	4	US-09-147-935A-46	Sequence 46, Appl
C 33	20	100.0	306	4	US-09-147-935A-47	Sequence 47, Appl
C 34	20	100.0	319	4	US-08-750-088A-35	Sequence 35, Appl
C 35	20	100.0	324	4	US-08-750-088A-36	Sequence 36, Appl
C 36	18.4	92.0	306	4	US-09-147-935A-20	Sequence 20, Appl
C 37	18.4	92.0	306	4	US-09-147-935A-21	Sequence 21, Appl
C 38	18.4	92.0	306	4	US-09-147-935A-23	Sequence 23, Appl
C 39	18.4	92.0	306	4	US-09-147-935A-25	Sequence 25, Appl
C 40	18.4	92.0	306	4	US-09-147-935A-26	Sequence 26, Appl
C 41	18.4	92.0	306	4	US-09-147-935A-34	Sequence 34, Appl
C 42	18.4	92.0	306	4	US-09-147-935A-42	Sequence 42, Appl
C 43	18.4	92.0	306	4	US-09-147-935A-44	Sequence 44, Appl
C 44	16.8	84.0	306	4	US-09-147-935A-2	Sequence 2, Appl
C 45	16.8	84.0	306	4	US-09-147-935A-6	Sequence 6, Appl

ALIGNMENTS

RESULT 1  
US-08-750-088A-69/c  
; Sequence 69, Application US/08750088A  
; Patent No. 6329138  
; GENERAL INFORMATION:  
; APPLICANT: DE BEENHOUWER, HANS  
; APPLICANT: PORTAELS, FRAN OISE  
; APPLICANT: MACHTELINCKX, LIEVE  
; APPLICANT: JANNES, GEERT  
; APPLICANT: ROSSAU, RUDI  
; TITLE OF INVENTION: METHOD FOR DETECTION OF THE ANTIBIOTIC  
; TITLE OF INVENTION: RESISTANCE SPECTRUM OF MYCOBACTERIUM SPECIES  
; NUMBER OF SEQUENCES: 71  
; CORRESPONDENCE ADDRESS:  
; ADDRESSED: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
; STREET: 1100 NEW YORK AVENUE, SUITE 600  
; CITY: WASHINGTON  
; STATE: D.C.  
; COUNTRY: US  
; ZIP: 20005-3934  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/750,088A  
; FILING DATE: 21-FEB-1997  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: GOLDSTEIN, JORGE A.  
; REGISTRATION NUMBER: 29,021  
; REFERENCE/DOCKET NUMBER: 1657.0010000  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-371-2600  
; TELEFAX: 202-371-2540  
; INFORMATION FOR SEQ ID NO: 69:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 228 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
US-08-750-088A-69

Query Match 100.0%; Score 20; DB 4; Length 228;  
Best Local Similarity 100.0%; Pred. No. 1.1;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATGTTGATCAGGTCCTGC 20

```
Db 88 GGATGTTGATCAGGCTCTGC 69
|||||
; SOFTWARE: KOPATIN 1.0
; SEQ ID NO 1
; LENGTH: 306
; TYPE: DNA
; ORGANISM: Mycobacterium abscessus
US-09-147-935A-1

Query Match 100.0%; Score 20; DB 4; Length 306;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATGTTGATCAGGCTCTGC 20
|||||
Db 129 GGATGTTGATCAGGCTCTGC 110

RESULT 4
US-09-147-935A-3/C
; Sequence 3, Application US/09147935A
; Patent No. 6242584
; GENERAL INFORMATION:
; APPLICANT: KOOK, Yoon-Hoh
; APPLICANT: KIM, Bum-Joon
; TITLE OF INVENTION: A METHOD FOR IDENTIFYING MYCOBACTERIAL SPECIES BY
; TITLE OF INVENTION: COMPARATIVE SEQUENCE ANALYSIS OF rpoB GENE
; FILE REFERENCE: 0136/0F425
; CURRENT APPLICATION NUMBER: US/09/147,935A
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: PCT/KR98/00228
; PRIOR FILING DATE: 1998-07-28
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: KOPATIN 1.0
; SEQ ID NO 3
; LENGTH: 306
; TYPE: DNA
; ORGANISM: Mycobacterium asiaticum
US-09-147-935A-3

Query Match 100.0%; Score 20; DB 4; Length 306;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATGTTGATCAGGCTCTGC 20
|||||
Db 129 GGATGTTGATCAGGCTCTGC 110

RESULT 5
US-09-147-935A-4/C
; Sequence 4, Application US/09147935A
; Patent No. 6242584
; GENERAL INFORMATION:
; APPLICANT: KOOK, Yoon-Hoh
; APPLICANT: KIM, Bum-Joon
; TITLE OF INVENTION: A METHOD FOR IDENTIFYING MYCOBACTERIAL SPECIES BY
; TITLE OF INVENTION: COMPARATIVE SEQUENCE ANALYSIS OF rpoB GENE
; FILE REFERENCE: 0136/0F425
; CURRENT APPLICATION NUMBER: US/09/147,935A
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: PCT/KR98/00228
; PRIOR FILING DATE: 1998-07-28
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: KOPATIN 1.0
; SEQ ID NO 4
; LENGTH: 306
; TYPE: DNA
; ORGANISM: Mycobacterium aurum
US-09-147-935A-4

Query Match 100.0%; Score 20; DB 4; Length 306;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

---

```
Db 88 GGATGTTGATCAGGCTCTGC 69
|||||
; SOFTWARE: KOPATIN 1.0
; SEQ ID NO 1
; LENGTH: 306
; TYPE: DNA
; ORGANISM: Mycobacterium abscessus
US-09-147-935A-1

Query Match 100.0%; Score 20; DB 4; Length 306;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATGTTGATCAGGCTCTGC 20
|||||
Db 129 GGATGTTGATCAGGCTCTGC 110

RESULT 4
US-09-147-935A-3/C
; Sequence 3, Application US/09147935A
; Patent No. 6242584
; GENERAL INFORMATION:
; APPLICANT: KOOK, Yoon-Hoh
; APPLICANT: KIM, Bum-Joon
; TITLE OF INVENTION: A METHOD FOR IDENTIFYING MYCOBACTERIAL SPECIES BY
; TITLE OF INVENTION: COMPARATIVE SEQUENCE ANALYSIS OF rpoB GENE
; FILE REFERENCE: 0136/0F425
; CURRENT APPLICATION NUMBER: US/09/147,935A
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: PCT/KR98/00228
; PRIOR FILING DATE: 1998-07-28
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: KOPATIN 1.0
; SEQ ID NO 3
; LENGTH: 306
; TYPE: DNA
; ORGANISM: Mycobacterium asiaticum
US-09-147-935A-3

Query Match 100.0%; Score 20; DB 4; Length 306;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATGTTGATCAGGCTCTGC 20
|||||
Db 129 GGATGTTGATCAGGCTCTGC 110

RESULT 5
US-09-147-935A-4/C
; Sequence 4, Application US/09147935A
; Patent No. 6242584
; GENERAL INFORMATION:
; APPLICANT: KOOK, Yoon-Hoh
; APPLICANT: KIM, Bum-Joon
; TITLE OF INVENTION: A METHOD FOR IDENTIFYING MYCOBACTERIAL SPECIES BY
; TITLE OF INVENTION: COMPARATIVE SEQUENCE ANALYSIS OF rpoB GENE
; FILE REFERENCE: 0136/0F425
; CURRENT APPLICATION NUMBER: US/09/147,935A
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: PCT/KR98/00228
; PRIOR FILING DATE: 1998-07-28
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: KOPATIN 1.0
; SEQ ID NO 4
; LENGTH: 306
; TYPE: DNA
; ORGANISM: Mycobacterium aurum
US-09-147-935A-4

Query Match 100.0%; Score 20; DB 4; Length 306;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

---

```
Db 88 GGATGTTGATCAGGCTCTGC 69
|||||
; SOFTWARE: KOPATIN 1.0
; SEQ ID NO 1
; LENGTH: 306
; TYPE: DNA
; ORGANISM: Mycobacterium abscessus
US-09-147-935A-1

Query Match 100.0%; Score 20; DB 4; Length 306;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATGTTGATCAGGCTCTGC 20
|||||
Db 129 GGATGTTGATCAGGCTCTGC 110

RESULT 4
US-09-147-935A-3/C
; Sequence 3, Application US/09147935A
; Patent No. 6242584
; GENERAL INFORMATION:
; APPLICANT: KOOK, Yoon-Hoh
; APPLICANT: KIM, Bum-Joon
; TITLE OF INVENTION: A METHOD FOR IDENTIFYING MYCOBACTERIAL SPECIES BY
; TITLE OF INVENTION: COMPARATIVE SEQUENCE ANALYSIS OF rpoB GENE
; FILE REFERENCE: 0136/0F425
; CURRENT APPLICATION NUMBER: US/09/147,935A
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: PCT/KR98/00228
; PRIOR FILING DATE: 1998-07-28
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: KOPATIN 1.0
; SEQ ID NO 3
; LENGTH: 306
; TYPE: DNA
; ORGANISM: Mycobacterium asiaticum
US-09-147-935A-3

Query Match 100.0%; Score 20; DB 4; Length 306;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATGTTGATCAGGCTCTGC 20
|||||
Db 129 GGATGTTGATCAGGCTCTGC 110

RESULT 5
US-09-147-935A-4/C
; Sequence 4, Application US/09147935A
; Patent No. 6242584
; GENERAL INFORMATION:
; APPLICANT: KOOK, Yoon-Hoh
; APPLICANT: KIM, Bum-Joon
; TITLE OF INVENTION: A METHOD FOR IDENTIFYING MYCOBACTERIAL SPECIES BY
; TITLE OF INVENTION: COMPARATIVE SEQUENCE ANALYSIS OF rpoB GENE
; FILE REFERENCE: 0136/0F425
; CURRENT APPLICATION NUMBER: US/09/147,935A
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: PCT/KR98/00228
; PRIOR FILING DATE: 1998-07-28
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: KOPATIN 1.0
; SEQ ID NO 4
; LENGTH: 306
; TYPE: DNA
; ORGANISM: Mycobacterium aurum
US-09-147-935A-4

Query Match 100.0%; Score 20; DB 4; Length 306;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```



QY 1 GGATGTTGATCAGGGTCTGC 20  
|||||  
Db 129 GGATGTTGATCAGGGTCTGC 110

RESULT 6  
US-09-147-935A-5/c  
; Sequence 5, Application US/09147935A  
; Patent No. 6242584  
; GENERAL INFORMATION:  
; APPLICANT: KOOK, Yoon-Hoh  
; APPLICANT: KIM, Bum-Joon  
; TITLE OF INVENTION: A METHOD FOR IDENTIFYING MYCOBACTERIAL SPECIES BY  
; FILE REFERENCE: 0136/OF425  
; CURRENT APPLICATION NUMBER: US/09/147,935A  
; CURRENT FILING DATE: 1999-03-19  
; PRIOR APPLICATION NUMBER: PCT/KR98/00228  
; PRIOR FILING DATE: 1998-07-28  
; NUMBER OF SEQ ID NOS: 50  
; SOFTWARE: KOPATIN 1.0  
; SEQ ID NO 5  
; LENGTH: 306  
; TYPE: DNA  
; ORGANISM: Mycobacterium avium  
US-09-147-935A-5

Query Match 100.0%; Score 20; DB 4; Length 306;  
Best Local Similarity 100.0%; Pred. No. 1.1;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATGTTGATCAGGGTCTGC 20  
|||||  
Db 129 GGATGTTGATCAGGGTCTGC 110

RESULT 7  
US-09-147-935A-8/c  
; Sequence 8, Application US/09147935A  
; Patent No. 6242584  
; GENERAL INFORMATION:  
; APPLICANT: KOOK, Yoon-Hoh  
; APPLICANT: KIM, Bum-Joon  
; TITLE OF INVENTION: A METHOD FOR IDENTIFYING MYCOBACTERIAL SPECIES BY  
; FILE REFERENCE: 0136/OF425  
; CURRENT APPLICATION NUMBER: US/09/147,935A  
; CURRENT FILING DATE: 1999-03-19  
; PRIOR APPLICATION NUMBER: PCT/KR98/00228  
; PRIOR FILING DATE: 1998-07-28  
; NUMBER OF SEQ ID NOS: 50  
; SOFTWARE: KOPATIN 1.0  
; SEQ ID NO 8  
; LENGTH: 306  
; TYPE: DNA  
; ORGANISM: Mycobacterium celatum Typel  
US-09-147-935A-8

Query Match 100.0%; Score 20; DB 4; Length 306;  
Best Local Similarity 100.0%; Pred. No. 1.1;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATGTTGATCAGGGTCTGC 20  
|||||  
Db 129 GGATGTTGATCAGGGTCTGC 110

RESULT 8  
US-09-147-935A-9/c  
; Sequence 9, Application US/09147935A  
; Patent No. 6242584  
; GENERAL INFORMATION:  
; APPLICANT: KOOK, Yoon-Hoh

; APPLICANT: KIM, Bum-Joon  
; TITLE OF INVENTION: A METHOD FOR IDENTIFYING MYCOBACTERIAL SPECIES BY  
; FILE REFERENCE: 0136/OF425  
; CURRENT APPLICATION NUMBER: US/09/147,935A  
; CURRENT FILING DATE: 1999-03-19  
; PRIOR APPLICATION NUMBER: PCT/KR98/00228  
; PRIOR FILING DATE: 1998-07-28  
; NUMBER OF SEQ ID NOS: 50  
; SOFTWARE: KOPATIN 1.0  
; SEQ ID NO 9  
; LENGTH: 306  
; TYPE: DNA  
; ORGANISM: Mycobacterium celatum Type2  
US-09-147-935A-9

Query Match 100.0%; Score 20; DB 4; Length 306;  
Best Local Similarity 100.0%; Pred. No. 1.1;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATGTTGATCAGGGTCTGC 20  
|||||  
Db 129 GGATGTTGATCAGGGTCTGC 110

RESULT 9  
US-09-147-935A-10/c  
; Sequence 10, Application US/09147935A  
; Patent No. 6242584  
; GENERAL INFORMATION:  
; APPLICANT: KOOK, Yoon-Hoh  
; APPLICANT: KIM, Bum-Joon  
; TITLE OF INVENTION: A METHOD FOR IDENTIFYING MYCOBACTERIAL SPECIES BY  
; FILE REFERENCE: 0136/OF425  
; CURRENT APPLICATION NUMBER: US/09/147,935A  
; CURRENT FILING DATE: 1999-03-19  
; PRIOR APPLICATION NUMBER: PCT/KR98/00228  
; PRIOR FILING DATE: 1998-07-28  
; NUMBER OF SEQ ID NOS: 50  
; SOFTWARE: KOPATIN 1.0  
; SEQ ID NO 10  
; LENGTH: 306  
; TYPE: DNA  
; ORGANISM: Mycobacterium chelonae  
US-09-147-935A-10

Query Match 100.0%; Score 20; DB 4; Length 306;  
Best Local Similarity 100.0%; Pred. No. 1.1;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATGTTGATCAGGGTCTGC 20  
|||||  
Db 129 GGATGTTGATCAGGGTCTGC 110

RESULT 10  
US-09-147-935A-11/c  
; Sequence 11, Application US/09147935A  
; Patent No. 6242584  
; GENERAL INFORMATION:  
; APPLICANT: KOOK, Yoon-Hoh  
; APPLICANT: KIM, Bum-Joon  
; TITLE OF INVENTION: A METHOD FOR IDENTIFYING MYCOBACTERIAL SPECIES BY  
; FILE REFERENCE: 0136/OF425  
; CURRENT APPLICATION NUMBER: US/09/147,935A  
; CURRENT FILING DATE: 1999-03-19  
; PRIOR APPLICATION NUMBER: PCT/KR98/00228  
; PRIOR FILING DATE: 1998-07-28  
; NUMBER OF SEQ ID NOS: 50  
; SOFTWARE: KOPATIN 1.0  
; SEQ ID NO 11

; LENGTH: 306  
; TYPE: DNA  
; ORGANISM: Mycobacterium chitae  
US-09-147-935A-11

Query Match 100.0%; Score 20; DB 4; Length 306;  
Best Local Similarity 100.0%; Pred. No. 1.1;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATGTTGATCAGGGTCTGC 20  
|||||

Db 129 GGATGTTGATCAGGGTCTGC 110

RESULT 11

US-09-147-935A-12/c  
; Sequence 12, Application US/09147935A

; Patent No. 6242584

; GENERAL INFORMATION:

; APPLICANT: KOOK, Yoon-Hoh

; TITLE OF INVENTION: A METHOD FOR IDENTIFYING MYCOBACTERIAL SPECIES BY

; FILE REFERENCE: 0136/0F425  
; CURRENT APPLICATION NUMBER: US/09/147,935A

; PRIOR FILING DATE: 1999-03-19

; PRIOR FILING DATE: 1998-07-28

; NUMBER OF SEQ ID NOS: 50

; SOFTWARE: KOPATIN 1.0

; SEQ ID NO 12

; LENGTH: 306

; TYPE: DNA

; ORGANISM: Mycobacterium fallax

US-09-147-935A-12

Query Match 100.0%; Score 20; DB 4; Length 306;  
Best Local Similarity 100.0%; Pred. No. 1.1;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATGTTGATCAGGGTCTGC 20  
|||||

Db 129 GGATGTTGATCAGGGTCTGC 110

RESULT 12

US-09-147-935A-13/c

; Sequence 13, Application US/09147935A

; Patent No. 6242584

; GENERAL INFORMATION:

; APPLICANT: KOOK, Yoon-Hoh

; TITLE OF INVENTION: A METHOD FOR IDENTIFYING MYCOBACTERIAL SPECIES BY

; FILE REFERENCE: 0136/0F425  
; CURRENT APPLICATION NUMBER: US/09/147,935A

; PRIOR FILING DATE: 1999-03-19

; PRIOR FILING DATE: 1998-07-28

; NUMBER OF SEQ ID NOS: 50

; SOFTWARE: KOPATIN 1.0

; SEQ ID NO 13

; LENGTH: 306

; TYPE: DNA

; ORGANISM: Mycobacterium flavescens

US-09-147-935A-13

Query Match 100.0%; Score 20; DB 4; Length 306;  
Best Local Similarity 100.0%; Pred. No. 1.1;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATGTTGATCAGGGTCTGC 20  
|||||

Db 129 GGATGTTGATCAGGGTCTGC 110

RESULT 13

US-09-147-935A-14/c

; Sequence 14, Application US/09147935A

; Patent No. 6242584

; GENERAL INFORMATION:

; APPLICANT: KOOK, Yoon-Hoh

; TITLE OF INVENTION: A METHOD FOR IDENTIFYING MYCOBACTERIAL SPECIES BY

; FILE REFERENCE: 0136/0F425  
; CURRENT APPLICATION NUMBER: US/09/147,935A

; PRIOR FILING DATE: 1999-03-19

; PRIOR FILING DATE: 1998-07-28

; NUMBER OF SEQ ID NOS: 50

; SOFTWARE: KOPATIN 1.0

; SEQ ID NO 14

; LENGTH: 306

; TYPE: DNA

; ORGANISM: Mycobacterium fortuitum

US-09-147-935A-14

Query Match 100.0%; Score 20; DB 4; Length 306;  
Best Local Similarity 100.0%; Pred. No. 1.1;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATGTTGATCAGGGTCTGC 20  
|||||

Db 129 GGATGTTGATCAGGGTCTGC 110

RESULT 14

US-09-147-935A-15/c

; Sequence 15, Application US/09147935A

; Patent No. 6242584

; GENERAL INFORMATION:

; APPLICANT: KOOK, Yoon-Hoh

; TITLE OF INVENTION: A METHOD FOR IDENTIFYING MYCOBACTERIAL SPECIES BY

; FILE REFERENCE: 0136/0F425  
; CURRENT APPLICATION NUMBER: US/09/147,935A

; PRIOR FILING DATE: 1999-03-19

; PRIOR FILING DATE: 1998-07-28

; NUMBER OF SEQ ID NOS: 50

; SOFTWARE: KOPATIN 1.0

; SEQ ID NO 15

; LENGTH: 306

; TYPE: DNA

; ORGANISM: Mycobacterium fortuitum

US-09-147-935A-15

Query Match 100.0%; Score 20; DB 4; Length 306;  
Best Local Similarity 100.0%; Pred. No. 1.1;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATGTTGATCAGGGTCTGC 20  
|||||

Db 129 GGATGTTGATCAGGGTCTGC 110

RESULT 15

US-09-147-935A-17/c

; Sequence 17, Application US/09147935A

; Patent No. 6242584

; GENERAL INFORMATION:

; APPLICANT: KOOK, Yoon-Hoh

; TITLE OF INVENTION: A METHOD FOR IDENTIFYING MYCOBACTERIAL SPECIES BY

; TITLE OF INVENTION: COMPARATIVE SEQUENCE ANALYSIS OF rpoB GENE  
; FILE REFERENCE: 0136/0F425  
; CURRENT APPLICATION NUMBER: US/09/147,935A  
; CURRENT FILING DATE: 1999-03-19  
; PRIOR APPLICATION NUMBER: PCT/RR98/00228  
; PRIOR FILING DATE: 1998-07-28  
; NUMBER OF SEQ ID NOS: 50  
; SOFTWARE: KOPATIN 1.0  
; SEQ ID NO 17  
; LENGTH: 306  
; TYPE: DNA  
; ORGANISM: Mycobacterium genavense  
US-09-147-935A-17

Query Match 100.0%; Score 20; DB 4; Length 306;  
Best Local Similarity 100.0%; Pred.No. 1.1;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 GGATGTTGATCAGGCTCTGC 20  
Db 129 GGATGTTGATCAGGCTCTGC 110  
|||||

Search completed: November 12, 2002, 21:03:47  
Job time : 3.68208 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 12, 2002, 15:44:35 : Search time 13.0983 Seconds  
(without alignments)  
3438.621 Million cell updates/sec

Title: US-09-697-123B-26  
Perfect score: 20  
Sequence: 1 ggaatgcatcagggtctgc 20

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_101002.\*  
1: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.\*  
2: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.\*  
3: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.\*  
4: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.\*  
5: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.\*  
6: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.\*  
7: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.\*  
8: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.\*  
9: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.\*  
10: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.\*  
11: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.\*  
12: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.\*  
13: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.\*  
14: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.\*  
15: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.\*  
16: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.\*  
17: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.\*  
18: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.\*  
19: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.\*  
20: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.\*  
21: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.\*  
22: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.\*  
23: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.\*  
24: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	100.0	20	22	AAS05226
C 2	20	100.0	228	17	AAT29617
C 3	20	100.0	254	17	AAT29615
C 4	20	100.0	300	24	AAS99529
C 5	20	100.0	306	19	AA27208
C 6	20	100.0	306	19	AA27209
C 7	20	100.0	306	19	AA27210
C 8	20	100.0	306	19	AA27211
C 9	20	100.0	306	19	AA27212

C 10	20	100.0	306	19	AA27213	RpoB gene fragment
C 11	20	100.0	306	19	AA27216	RpoB gene fragment
C 12	20	100.0	306	19	AA27219	RpoB gene fragment
C 13	20	100.0	306	19	AA27220	RpoB gene fragment
C 14	20	100.0	306	19	AA27221	RpoB gene fragment
C 15	20	100.0	306	19	AA27222	RpoB gene fragment
C 16	20	100.0	306	19	AA27223	RpoB gene fragment
C 17	20	100.0	306	19	AA27224	RpoB gene fragment
C 18	20	100.0	306	19	AA27225	RpoB gene fragment
C 19	20	100.0	306	19	AA27226	RpoB gene fragment
C 20	20	100.0	306	19	AA27227	RpoB gene fragment
C 21	20	100.0	306	19	AA27228	RpoB gene fragment
C 22	20	100.0	306	19	AA27229	RpoB gene fragment
C 23	20	100.0	306	19	AA27230	RpoB gene fragment
C 24	20	100.0	306	19	AA27231	RpoB gene fragment
C 25	20	100.0	306	19	AA27232	RpoB gene fragment
C 26	20	100.0	306	19	AA27233	RpoB gene fragment
C 27	20	100.0	306	19	AA27234	RpoB gene fragment
C 28	20	100.0	306	19	AA27235	RpoB gene fragment
C 29	20	100.0	306	19	AA27236	RpoB gene fragment
C 30	20	100.0	306	19	AA27237	RpoB gene fragment
C 31	20	100.0	306	19	AA27238	RpoB gene fragment
C 32	20	100.0	306	19	AA27239	RpoB gene fragment
C 33	20	100.0	306	19	AA27240	RpoB gene fragment
C 34	20	100.0	306	19	AA27241	RpoB gene fragment
C 35	20	100.0	306	19	AA27242	RpoB gene fragment
C 36	20	100.0	306	24	AAS99528	Mycobacterium spec
C 37	20	100.0	306	24	AAS99532	Mycobacterium spec
C 38	20	100.0	306	24	AAS99533	Mycobacterium spec
C 39	20	100.0	306	24	AAS99535	Mycobacterium spec
C 40	20	100.0	306	24	AAS99536	Mycobacterium spec
C 41	20	100.0	306	24	AAS99537	Mycobacterium spec
C 42	20	100.0	306	24	AAS99540	Mycobacterium spec
C 43	20	100.0	306	24	AAS99541	Mycobacterium spec
C 44	20	100.0	306	24	AAS99545	Mycobacterium spec
C 45	20	100.0	306	24	AAS99546	Mycobacterium spec

ALIGNMENTS

RESULT 1  
AAS05226  
ID AAS05226 standard; DNA; 20 Bp.

XX AAS05226;

XX 07-SEP-2001 (first entry)

XX Mycobacterium species rpoB gene fragment PCR primer RPO3'.

XX Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;  
XX PCR-restriction fragment length polymorphism analysis; PCR primer; ss.

OS Mycobacterium sp.

PN WO200131061-A1.

XX 03-MAY-2001.

XX 27-OCT-2000; 2000WO-KR01223.

XX 27-OCT-1999; 99KR-0046795.

XX (ERUM-) ERUME BIOTECH CO LTD.

XX Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;

XX WPI; 2001-300520/31.

XX New DNA fragments from the rpoB gene of mycobacteria, useful for  
PT diagnosis and identification of many mycobacterial species by  
PT restriction fragment length polymorphism.

XX PS Disclosure; Page 13; 50pp; English.

XX CC The present sequence for Mycobacterium species rpoB gene fragment PCR

XX CC primer RPO3' is used with PCR primer RPO5' (AAS05225) to amplify the

XX CC rpoB gene region from 24 rpoB gene fragments (AAS05201-AAS05224) from

XX CC various Mycobacterial species. These rpoB gene fragments can be used

XX CC in the diagnosis and identification of Mycobacterium species using a

XX CC novel PCR-restriction fragment length polymorphism analysis (PRA)

XX CC method. The method comprises obtaining a restriction fragment length

XX CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,

XX CC amplifying and digesting the DNA fragment from the microorganism to

XX CC be identified and comparing the RFLP patterns from the known rpoB gene

XX CC fragments with the unidentified fragment. The rpoB gene fragments

XX CC are useful to identify a wide range of Mycobacterium species, e.g. for

XX CC diagnosis or to obtain epidemiological and pathogenesis information for

XX CC selection of appropriate therapies including M. tuberculosis, M. leprae

XX CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected

XX CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene

XX CC fragments is rapid, precise, simple and cost effective (only 1 PCR

XX CC required), and can differentiate between many species in a single

XX CC experiment, including those difficult to distinguish by usual biochemical

XX CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for

XX CC detecting specific Mycobacterial species.

SQ Sequence 20 BP; 3 A; 3 C; 8 G; 6 T; 0 other;

Query Match 100.0%; Score 20; DB 22; Length 20;  
Best Local Similarity 100.0%; Pred. No. 2.3;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATGTTGATCAGGGTCTGC 20  
|||||

DB 1 GGATGTTGATCAGGGTCTGC 20

# RESULT 2

AAT29617/C  
ID AAT29617 standard; DNA; 228 BP.

XX AC AAT29617;

XX DT 11-JUL-1996 (first entry)

XX DE Partial sequence of Mycobacterium MAC strain ITG 926 rpoB gene.

XX KW Antibiotic; resistance; spectrum; gene; mycobacterium;  
KW determination; amplification; MAC strain; rpoB; fragment;  
KW probe; differential; hybridisation; pattern; rifampicin;  
KW rifabutin; species identification; ss.

XX OS Mycobacterium MAC strain.

XX PN W09533851-A2.

XX PD 14-DEC-1995.

XX PF 09-JUN-1995; 95WO-EP02230.

XX PR 09-JUN-1994; 94EP-0870093.

XX PA (INNO-) INNOGENETICS NV.

XX PI De Beenhouwer H, Jannes G, Machtelinckx L, Portael F;  
PI Rossau R;

XX DR WPI; 1996-040250/04.

XX PT Probes and primers for determin. of antibiotic resistance spectrum of  
PT Mycobacterium, opt. coupled with species identification - from  
PT different patterns of hybridisation with rpoB gene

XX PS Claim 13; Fig 11; 69pp; English.

XX CC The antibiotic resistance spectrum (ARS) of a mycobacterium can be  
CC determined by amplifying the relevant part of the antibiotic  
CC resistance gene, hybridising it with at least 1 rpoB gene probe,  
CC detecting the hybrids formed and inferring the ARS, and opt. the  
CC spp. using species specific probes (i.e. AAT12148/49 derived from  
CC the present sequence, the partial nucleotide sequence of the  
CC presumptive Mycobacterium MAC strain ITG 926 rpoB gene) from the  
CC differential hybridisation patterns. The method is partic. useful  
CC for the detection of rifampicin and/or rifabutin resistance in  
CC M. leprae or M. tuberculosis, and mycobacterial spp.  
CC identification. The method is rapid and reliable and provides  
CC simultaneous determin. of ARS and spp. identity.

SQ Sequence 228 BP; 40 A; 76 C; 73 G; 37 T; 2 other;

Query Match 100.0%; Score 20; DB 17; Length 228;  
Best Local Similarity 100.0%; Pred. No. 3.1;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATGTTGATCAGGGTCTGC 20  
|||||

DB 88 GGATGTTGATCAGGGTCTGC 69

# RESULT 3

AAT29615/C

ID AAT29615 standard; DNA; 254 BP.

XX AC AAT29615;

XX DT 11-JUL-1996 (first entry)

XX DE Partial sequence of M. scrofulaceum strain ITG 4979 rpoB gene.

XX KW Antibiotic; resistance; spectrum; gene; mycobacterium;  
KW determination; amplification; scrofulaceum; rpoB; fragment;  
KW probe; differential; hybridisation; pattern; rifampicin;  
KW rifabutin; species identification; ss.

XX OS Mycobacterium scrofulaceum.

XX PN W09533851-A2.

XX PD 14-DEC-1995.

XX PF 09-JUN-1995; 95WO-EP02230.

XX PR 09-JUN-1994; 94EP-0870093.

XX PA (INNO-) INNOGENETICS NV.

XX PI De Beenhouwer H, Jannes G, Machtelinckx L, Portael F;  
PI Rossau R;

XX DR WPI; 1996-040250/04.

XX PT Probes and primers for determin. of antibiotic resistance spectrum of  
PT Mycobacterium, opt. coupled with species identification - from  
PT different patterns of hybridisation with rpoB gene

XX PS Claim 13; Fig 7; 69pp; English.

XX CC The antibiotic resistance spectrum (ARS) of a mycobacterium can be  
CC determined by amplifying the relevant part of the antibiotic  
CC resistance gene, hybridising it with at least 1 rpoB gene probe,  
CC detecting the hybrids formed and inferring the ARS, and opt. the  
CC spp. using species specific probes (i.e. AAT12146 derived from the  
CC present sequence, the partial nucleotide sequence of the  
CC presumptive M. scrofulaceum strain ITG 4979 rpoB gene) from the  
CC differential hybridisation patterns. The method is partic. useful  
CC for the detection of rifampicin and/or rifabutin resistance in  
CC M. leprae or M. tuberculosis, and mycobacterial spp.

CC Identification. The method is rapid and reliable and provides  
 CC simultaneous determin. of ARS and spp. identity.

XX Sequence 254 BP; 40 A; 84 C; 90 G; 39 T; 1 other;

Query Match 100.0%; Score 20; DB 17; Length 254;  
 Best Local Similarity 100.0%; Pred. No. 3.1; Mismatches 0; Indels 0; Gaps 0;  
 Matches 20; Conservative 0;

OY 1 GGATGTTGATCAGGGTCTGC 20  
 |||||  
 DB 94 GGATGTTGATCAGGGTCTGC 75

## RESULT 4

AAS99529/c  
 ID AAS99529 standard; DNA; 300 BP.

XX AC AAS99529;

XX DT 12-MAR-2002 (first entry)

DE Mycobacterium species identification primer #4.

KW Drug resistance detection; mycobacterial species identification; probe;  
 KW oligonucleotide chip; rpoB; sputum; blood; cerebrospinal fluid; ss;  
 KW primer.

OS Mycobacterium avium.

PN WO200192573-A1.

XX PD 06-DEC-2001.

XX PF 30-MAY-2001; 2001WO-KR00904.

XX PR 30-MAY-2000; 2000KR-0029369.

XX PA (BIOM-) BIOMEDLAB CO LTD.

PI Kim H, Kim N, Yoon S, Kim J, Park M;

DR WPI; 2002-075472/10.

PT Kit for mycobacterial species identification and drug resistance  
 PT detection, has oligonucleotide chip with species identification probe,  
 PT a mycobacterial drug-resistance detection probe, and its contrast group  
 PT probe.

PS Disclosure; Page 21; 74pp; English.

CC The invention relates to a diagnostic kit for mycobacterial species  
 CC identification and drug resistance detection comprising an  
 CC oligonucleotide chip including a species identification probe, a  
 CC mycobacterial drug-resistance detection probe, a contrast group probe  
 CC corresponding to each drug resistance detection probe, and a marker for  
 CC detecting a hybridisation of the oligonucleotide chip and a specimen. The  
 CC identification probe is comprised of species-specific DNA sequences of  
 CC mycobacterial rpoB gene and the detection probe is comprised of one or  
 CC more modified codons of mycobacterial rpoB gene. The method involves  
 CC amplifying rpoB gene fragments of specimen by Polymerase Chain Reaction  
 CC (PCR) and discriminating species by fluorescent intensity corresponding  
 CC to a particular species. The specimen is preferably uncultured sputum,  
 CC blood or cerebrospinal fluid of a patient. Sequences AAS99478-AAS99569  
 CC represent mycobacterium species identification probes and primers of the  
 CC invention.

XX Sequence 300 BP; 49 A; 105 C; 101 G; 45 T; 0 other;

Query Match 100.0%; Score 20; DB 24; Length 300;  
 Best Local Similarity 100.0%; Pred. No. 3.2;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGATGTTGATCAGGGTCTGC 20  
 |||||  
 DB 129 GGATGTTGATCAGGGTCTGC 110

## RESULT 5

AAX27208/c

ID AAX27208 standard; DNA; 306 BP.

XX AC AAX27208;

XX DT 27-MAY-1999 (first entry)

XX DE RpoB gene fragment.

XX RpoB gene; mycobacteria; phylogenetic tree construction;

KW mycobacterial species identification; phylogenetic analysis; ss.

XX OS Mycobacteria simiae.

XX PN WO9905316-A1.

XX PD 04-FEB-1999.

XX PF 28-JUL-1998; 98KR-0000228.

XX PR 28-JUL-1997; 97KR-0035501.

XX PA (BION-) BIONEER CORP.

XX PI Kim B, Kook Y;

XX DR WPI; 1998-539367/46.

XX New pair of polymerase chain reaction (PCR) primers - for  
 PT sequence-specific amplification of the rpoB gene from mycobacterial  
 PT species, useful for detecting and identifying mycobacterial species

XX Claim 37; Page 73-74; 91pp; English.

CC This sequence represents a mycobacterial rpoB gene fragment, that is  
 CC amplified using the PCR primers of the invention. The primers form a  
 CC method of detecting and identifying mycobacterial species by constructing  
 CC a phylogenetic tree for the species. The use of the primers for  
 CC sequence-specific amplification of the rpoB gene (encoding the beta  
 CC subunit of RNA polymerase) from mycobacterial species provides an  
 CC efficient way of characterising these species. In addition to  
 CC phylogenetic analysis, the rpoB gene can be used as an alternative to  
 CC the 16S rRNA gene because it has four subunits, which are highly  
 CC conserved throughout prokaryotes. The method is particularly useful for  
 CC slow growing, fastidious or uncultivable mycobacteria. Also, rifampin  
 CC susceptibility can be simultaneously determined in M. tuberculosis.

XX Sequence 306 BP; 53 A; 101 C; 100 G; 52 T; 0 other;

Query Match 100.0%; Score 20; DB 19; Length 306;  
 Best Local Similarity 100.0%; Pred. No. 3.2;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGATGTTGATCAGGGTCTGC 20  
 |||||  
 DB 129 GGATGTTGATCAGGGTCTGC 110

## RESULT 6

AAX27209/c

ID AAX27209 standard; DNA; 306 BP.

XX AC AAX27209;

XX DT 27-MAY-1999 (first entry)

XX DE RpoB gene fragment.

```

XX RpoB gene; mycobacteria; phylogenetic tree construction;
KW mycobacterial species identification; phylogenetic analysis; ss.
XX
OS Mycobacteria smegmatis.
XX
PN WO9905316-A1.
XX
PD 04-FEB-1999.
XX
XX 28-JUL-1998; 98KR-0000228.
PF
XX
PR 28-JUL-1997; 97KR-0035501.
XX
XX (BION-) BIONEER CORP.
PA
XX Kim B, Kook Y;
XX
XX WPI; 1998-539367/46.
XX
XX New pair of polymerase chain reaction (PCR) primers - for
PT sequence-specific amplification of the rpoB gene from mycobacterial
PT species, useful for detecting and identifying mycobacterial species
XX
XX Claim 38; Page 74; 91pp; English.
XX
XX This sequence represents a mycobacterial rpoB gene fragment, that is
CC amplified using the PCR primers of the invention. The primers form a
CC method of detecting and identifying mycobacterial species by constructing
CC a phylogenetic tree for the species. The use of the primers for
CC sequence-specific amplification of the rpoB gene (encoding the beta
CC subunit of RNA polymerase) from mycobacterial species provides an
CC efficient way of characterising these species. In addition to
CC phylogenetic analysis, the rpoB gene can be used as an alternative to
CC the 16S rRNA gene because it has four subunits, which are highly
CC conserved throughout prokaryotes. The method is particularly useful for
CC slow growing, fastidious or uncultivable mycobacteria. Also, rifampin
CC susceptibility can be simultaneously determined in M. tuberculosis.
XX
XX Sequence 306 BP; 50 A; 108 C; 96 G; 52 T; 0 other;
SQ
Query Match 100.0%; Score 20; DB 19; Length 306;
Best Local Similarity 100.0%; Pred. NO. 3.2;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GGATGTTGATCAGGCTGTCG 20
Db 129 GGATGTTGATCAGGCTGTCG 110
RESULT 7
AAAX27210/C
ID AAAX27210 standard; DNA; 306 BP.
XX
XX AC AAAX27210;
XX
XX DT 27-MAY-1999 (first entry)
XX
XX DE RpoB gene fragment.
XX
XX KW RpoB gene; mycobacteria; phylogenetic tree construction;
KW mycobacterial species identification; phylogenetic analysis; ss.
XX
XX OS Mycobacteria szulgai.
XX
XX PN WO9905316-A1.
XX
XX PD 04-FEB-1999.
XX
XX PF 28-JUL-1998; 98KR-0000228.
XX
XX PR 28-JUL-1997; 97KR-0035501.
XX

```

```

PA (BION-) BIONEER CORP.
XX
XX Kim B, Kook Y;
XX
XX WPI; 1998-539367/46.
XX
XX New pair of polymerase chain reaction (PCR) primers - for
PT sequence-specific amplification of the rpoB gene from mycobacterial
PT species, useful for detecting and identifying mycobacterial species
XX
XX Claim 39; Page 74; 91pp; English.
XX
XX This sequence represents a mycobacterial rpoB gene fragment, that is
CC amplified using the PCR primers of the invention. The primers form a
CC method of detecting and identifying mycobacterial species by constructing
CC a phylogenetic tree for the species. The use of the primers for
CC sequence-specific amplification of the rpoB gene (encoding the beta
CC subunit of RNA polymerase) from mycobacterial species provides an
CC efficient way of characterising these species. In addition to
CC phylogenetic analysis, the rpoB gene can be used as an alternative to
CC the 16S rRNA gene because it has four subunits, which are highly
CC conserved throughout prokaryotes. The method is particularly useful for
CC slow growing, fastidious or uncultivable mycobacteria. Also, rifampin
CC susceptibility can be simultaneously determined in M. tuberculosis.
XX
XX Sequence 306 BP; 50 A; 104 C; 102 G; 50 T; 0 other;
SQ
Query Match 100.0%; Score 20; DB 19; Length 306;
Best Local Similarity 100.0%; Pred. NO. 3.2;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GGATGTTGATCAGGCTGTCG 20
Db 129 GGATGTTGATCAGGCTGTCG 110
RESULT 8
AAAX27211/C
ID AAAX27211 standard; DNA; 306 BP.
XX
XX AC AAAX27211;
XX
XX DT 27-MAY-1999 (first entry)
XX
XX DE RpoB gene fragment.
XX
XX KW RpoB gene; mycobacteria; phylogenetic tree construction;
KW mycobacterial species identification; phylogenetic analysis; ss.
XX
XX OS Mycobacteria terrae.
XX
XX PN WO9905316-A1.
XX
XX PD 04-FEB-1999.
XX
XX PF 28-JUL-1998; 98KR-0000228.
XX
XX PR 28-JUL-1997; 97KR-0035501.
XX
XX PA (BION-) BIONEER CORP.
XX
XX Kim B, Kook Y;
XX
XX WPI; 1998-539367/46.
XX
XX New pair of polymerase chain reaction (PCR) primers - for
PT sequence-specific amplification of the rpoB gene from mycobacterial
PT species, useful for detecting and identifying mycobacterial species
XX
XX Claim 40; Page 74-75; 91pp; English.
XX
XX This sequence represents a mycobacterial rpoB gene fragment, that is
CC amplified using the PCR primers of the invention. The primers form a

```

CC method of detecting and identifying mycobacterial species by constructing  
 CC a phylogenetic tree for the species. The use of the primers for  
 CC sequence-specific amplification of the rpoB gene (encoding the beta  
 CC subunit of RNA polymerase) from mycobacterial species provides an  
 CC efficient way of characterising these species. In addition to  
 CC phylogenetic analysis, the rpoB gene can be used as an alternative to  
 CC the 16S rRNA gene because it has four subunits, which are highly  
 CC conserved throughout prokaryotes. The method is particularly useful for  
 CC slow growing, fastidious or uncultivable mycobacteria. Also, rifampin  
 CC susceptibility can be simultaneously determined in *M. tuberculosis*.  
 XX  
 SQ Sequence 306 BP; 49 A; 102 C; 103 G; 52 T; 0 other;

Query Match 100.0%; Score 20; DB 19; Length 306;  
 Best Local Similarity 100.0%; Pred. No. 3.2;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGATGTTGATCAGGGTCTGC 20  
 Db 129 GGATGTTGATCAGGGTCTGC 110

## RESULT 9

AAAX27212/C  
 ID AAX27212 standard; DNA; 306 BP.

XX AC AAX27212;

XX DT 27-MAY-1999 (first entry)

XX DE RpoB gene fragment.

XX KW RpoB gene; mycobacteria; phylogenetic tree construction;  
 XX KW mycobacterial species identification; phylogenetic analysis; ss.

XX OS Mycobacteria thermoresistable.

XX PN WO9905316-A1.

XX PD 04-FEB-1999.

XX PF 28-JUL-1998; 98KR-0000228.

XX PR 28-JUL-1997; 97KR-0035501.

XX PA (BION-) BIONEER CORP.

XX PI Kim B, Kook Y;

XX DR WPI; 1998-539367/46.

XX PT New pair of polymerase chain reaction (PCR) primers - for  
 PT sequence-specific amplification of the rpoB gene from mycobacterial  
 PT species, useful for detecting and identifying mycobacterial species

XX PS Claim 41; Page 75; 91pp; English.

XX This sequence represents a mycobacterial rpoB gene fragment, that is  
 CC amplified using the PCR primers of the invention. The primers form a  
 CC method of detecting and identifying mycobacterial species by constructing  
 CC a phylogenetic tree for the species. The use of the primers for  
 CC sequence-specific amplification of the rpoB gene (encoding the beta  
 CC subunit of RNA polymerase) from mycobacterial species provides an  
 CC efficient way of characterising these species. In addition to  
 CC phylogenetic analysis, the rpoB gene can be used as an alternative to  
 CC the 16S rRNA gene because it has four subunits, which are highly  
 CC conserved throughout prokaryotes. The method is particularly useful for  
 CC slow growing, fastidious or uncultivable mycobacteria. Also, rifampin  
 CC susceptibility can be simultaneously determined in *M. tuberculosis*.  
 XX  
 SQ Sequence 306 BP; 51 A; 111 C; 102 G; 42 T; 0 other;

Query Match 100.0%; Score 20; DB 19; Length 306;

Best Local Similarity 100.0%; Pred. No. 3.2;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGATGTTGATCAGGGTCTGC 20  
 Db 129 GGATGTTGATCAGGGTCTGC 110

## RESULT 10

AAAX27213/C

ID AAX27213 standard; DNA; 306 BP.

XX AC AAX27213;

XX DT 27-MAY-1999 (first entry)

XX DE RpoB gene fragment.

XX KW RpoB gene; mycobacteria; phylogenetic tree construction;  
 XX KW mycobacterial species identification; phylogenetic analysis; ss.

XX OS Mycobacteria triviale.

XX PN WO9905316-A1.

XX PD 04-FEB-1999.

XX PF 28-JUL-1998; 98KR-0000228.

XX PR 28-JUL-1997; 97KR-0035501.

XX PA (BION-) BIONEER CORP.

XX PI Kim B, Kook Y;

XX DR WPI; 1998-539367/46.

XX PT New pair of polymerase chain reaction (PCR) primers - for  
 PT sequence-specific amplification of the rpoB gene from mycobacterial  
 PT species, useful for detecting and identifying mycobacterial species

XX PS Claim 42; Page 75; 91pp; English.

XX This sequence represents a mycobacterial rpoB gene fragment, that is  
 CC amplified using the PCR primers of the invention. The primers form a  
 CC method of detecting and identifying mycobacterial species by constructing  
 CC a phylogenetic tree for the species. The use of the primers for  
 CC sequence-specific amplification of the rpoB gene (encoding the beta  
 CC subunit of RNA polymerase) from mycobacterial species provides an  
 CC efficient way of characterising these species. In addition to  
 CC phylogenetic analysis, the rpoB gene can be used as an alternative to  
 CC the 16S rRNA gene because it has four subunits, which are highly  
 CC conserved throughout prokaryotes. The method is particularly useful for  
 CC slow growing, fastidious or uncultivable mycobacteria. Also, rifampin  
 CC susceptibility can be simultaneously determined in *M. tuberculosis*.  
 XX  
 SQ Sequence 306 BP; 50 A; 104 C; 109 G; 43 T; 0 other;

Query Match 100.0%; Score 20; DB 19; Length 306;

Best Local Similarity 100.0%; Pred. No. 3.2;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGATGTTGATCAGGGTCTGC 20  
 Db 129 GGATGTTGATCAGGGTCTGC 110

## RESULT 11

AAAX27216/C

ID AAX27216 standard; DNA; 306 BP.

XX AC AAX27216;

XX XX



```

DT 27-MAY-1999 (first entry)
DE RpoB gene fragment.
XX
KW RpoB gene; mycobacteria; phylogenetic tree construction;
KW mycobacterial species identification; phylogenetic analysis; ss.
OS Mycobacteria vaccae.
XX
PN WO9905316-A1.
XX
PD 04-FEB-1999.
XX
PF 28-JUL-1998; 98KR-0000228.
XX
PR 28-JUL-1997; 97KR-0035501.
XX
PA (BION-) BIONEER CORP.
XX
PI Kim B, Kook Y;
XX
XX WPI; 1998-539367/46.
XX
PT New pair of polymerase chain reaction (PCR) primers - for
PT sequence-specific amplification of the rpoB gene from mycobacterial
PT species, useful for detecting and identifying mycobacterial species
XX
PS Claim 45; Page 76; 91pp; English.
XX
CC This sequence represents a mycobacterial rpoB gene fragment, that is
CC amplified using the PCR primers of the invention. The primers form a
CC method of detecting and identifying mycobacterial species by constructing
CC a phylogenetic tree for the species. The use of the primers for
CC sequence-specific amplification of the rpoB gene (encoding the beta
CC subunit of RNA polymerase) from mycobacterial species provides an
CC efficient way of characterising these species. In addition to
CC phylogenetic analysis, the rpoB gene can be used as an alternative to
CC the 16S rRNA gene because it has four subunits, which are highly
CC conserved throughout prokaryotes. The method is particularly useful for
CC slow growing, fastidious or uncultivable mycobacteria. Also, rifampin
CC susceptibility can be simultaneously determined in M. tuberculosis.
XX
SQ Sequence 306 BP; 51 A; 109 C; 95 G; 51 T; 0 other;

Query Match 100.0%; Score 20; DB 19; Length 306;
Best Local Similarity 100.0%; Pred. No. 3.2;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATGTTGATCAGGGTCTGC 20
    |||||
Db 129 GGATGTTGATCAGGGTCTGC 110

RESULT 12
AAX27219/c
ID AAX27219 standard; DNA; 306 BP.
XX
AC AAX27219;
XX
XX 27-MAY-1999 (first entry)
DE RpoB gene fragment.
XX
KW RpoB gene; mycobacteria; phylogenetic tree construction;
KW mycobacterial species identification; phylogenetic analysis; ss.
OS Nocardia nova.
XX
PN WO9905316-A1.
XX
PD 04-FEB-1999.
XX
PF 28-JUL-1998; 98KR-0000228.

```

```

XX
PR 28-JUL-1997; 97KR-0035501.
XX
PA (BION-) BIONEER CORP.
XX
PI Kim B, Kook Y;
XX
XX WPI; 1998-539367/46.
XX
PT New pair of polymerase chain reaction (PCR) primers - for
PT sequence-specific amplification of the rpoB gene from mycobacterial
PT species, useful for detecting and identifying mycobacterial species
XX
PS Claim 48; Page 77; 91pp; English.
XX
CC This sequence represents a mycobacterial rpoB gene fragment, that is
CC amplified using the PCR primers of the invention. The primers form a
CC method of detecting and identifying mycobacterial species by constructing
CC a phylogenetic tree for the species. The use of the primers for
CC sequence-specific amplification of the rpoB gene (encoding the beta
CC subunit of RNA polymerase) from mycobacterial species provides an
CC efficient way of characterising these species. In addition to
CC phylogenetic analysis, the rpoB gene can be used as an alternative to
CC the 16S rRNA gene because it has four subunits, which are highly
CC conserved throughout prokaryotes. The method is particularly useful for
CC slow growing, fastidious or uncultivable mycobacteria. Also, rifampin
CC susceptibility can be simultaneously determined in M. tuberculosis.
XX
SQ Sequence 306 BP; 53 A; 109 C; 94 G; 50 T; 0 other;

Query Match 100.0%; Score 20; DB 19; Length 306;
Best Local Similarity 100.0%; Pred. No. 3.2;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATGTTGATCAGGGTCTGC 20
    |||||
Db 129 GGATGTTGATCAGGGTCTGC 110

RESULT 13
AAX27220/c
ID AAX27220 standard; DNA; 306 BP.
XX
AC AAX27220;
XX
XX 27-MAY-1999 (first entry)
DE RpoB gene fragment.
XX
KW RpoB gene; mycobacteria; phylogenetic tree construction;
KW mycobacterial species identification; phylogenetic analysis; ss.
OS Rhodococcus equi.
XX
PN WO9905316-A1.
XX
XX 04-FEB-1999.
XX
PF 28-JUL-1998; 98KR-0000228.
XX
PR 28-JUL-1997; 97KR-0035501.
XX
PA (BION-) BIONEER CORP.
XX
PI Kim B, Kook Y;
XX
XX WPI; 1998-539367/46.
XX
PT New pair of polymerase chain reaction (PCR) primers - for
PT sequence-specific amplification of the rpoB gene from mycobacterial
PT species, useful for detecting and identifying mycobacterial species
XX
PS Claim 49; Page 77-78; 91pp; English.

```

XX This sequence represents a mycobacterial rpoB gene fragment, that is  
 CC amplified using the PCR primers of the invention. The primers form a  
 CC method of detecting and identifying mycobacterial species by constructing  
 CC a phylogenetic tree for the species. The use of the primers for  
 CC sequence-specific amplification of the rpoB gene (encoding the beta  
 CC subunit of RNA polymerase) from mycobacterial species provides an  
 CC efficient way of characterising these species. In addition to  
 CC phylogenetic analysis, the rpoB gene can be used as an alternative to  
 CC the 16S rRNA gene because it has four subunits, which are highly  
 CC conserved throughout prokaryotes. The method is particularly useful for  
 CC slow growing, fastidious or uncultivable mycobacteria. Also, rifampin  
 CC susceptibility can be simultaneously determined in M. tuberculosis.  
 XX  
 SQ Sequence 306 BP; 50 A; 110 C; 98 G; 48 T; 0 other;

Query Match 100.0%; Score 20; DB 19; Length 306;  
 Best Local Similarity 100.0%; Pred. No. 3.2;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATGTTGATCAGGGTCTGC 20  
 |||||  
 Db 129 GGATGTTGATCAGGGTCTGC 110

RESULT 14  
 AAX27191/c  
 ID AAX27191 standard; DNA; 306 BP.  
 XX  
 AC AAX27191;  
 XX  
 DT 27-MAY-1999 (first entry)  
 XX  
 DE RpoB gene fragment.  
 XX  
 KW RpoB gene; mycobacteria; phylogenetic tree construction;  
 KW mycobacterial species identification; phylogenetic analysis; ss.  
 XX  
 OS Mycobacteria gordoniae.  
 XX  
 PN WO9905316-A1.  
 XX  
 PD 04-FEB-1999.  
 XX  
 PF 28-JUL-1998; 98KR-0000228.  
 XX  
 PR 28-JUL-1997; 97KR-0035501.  
 XX  
 PA (BION-) BIONEER CORP.  
 XX  
 PI Kim B, Kook Y;  
 XX  
 WP1; 1998-539367/46.

XX New pair of polymerase chain reaction (PCR) primers - for  
 PT sequence-specific amplification of the rpoB gene from mycobacterial  
 PT species, useful for detecting and identifying mycobacterial species  
 XX  
 PS Claim 20; Page 68; 91pp; English.  
 XX  
 CC This sequence represents a mycobacterial rpoB gene fragment, that is  
 CC amplified using the PCR primers of the invention. The primers form a  
 CC method of detecting and identifying mycobacterial species by constructing  
 CC a phylogenetic tree for the species. The use of the primers for  
 CC sequence-specific amplification of the rpoB gene (encoding the beta  
 CC subunit of RNA polymerase) from mycobacterial species provides an  
 CC efficient way of characterising these species. In addition to  
 CC phylogenetic analysis, the rpoB gene can be used as an alternative to  
 CC the 16S rRNA gene because it has four subunits, which are highly  
 CC conserved throughout prokaryotes. The method is particularly useful for  
 CC slow growing, fastidious or uncultivable mycobacteria. Also, rifampin  
 CC susceptibility can be simultaneously determined in M. tuberculosis.  
 XX

XX This sequence represents a mycobacterial rpoB gene fragment, that is  
 CC amplified using the PCR primers of the invention. The primers form a  
 CC method of detecting and identifying mycobacterial species by constructing  
 CC a phylogenetic tree for the species. The use of the primers for  
 CC sequence-specific amplification of the rpoB gene (encoding the beta  
 CC subunit of RNA polymerase) from mycobacterial species provides an  
 CC efficient way of characterising these species. In addition to  
 CC phylogenetic analysis, the rpoB gene can be used as an alternative to  
 CC the 16S rRNA gene because it has four subunits, which are highly  
 CC conserved throughout prokaryotes. The method is particularly useful for  
 CC slow growing, fastidious or uncultivable mycobacteria. Also, rifampin  
 CC susceptibility can be simultaneously determined in M. tuberculosis.  
 XX

SQ Sequence 306 BP; 52 A; 101 C; 101 G; 52 T; 0 other;

Query Match 100.0%; Score 20; DB 19; Length 306;  
 Best Local Similarity 100.0%; Pred. No. 3.2;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATGTTGATCAGGGTCTGC 20  
 |||||  
 Db 129 GGATGTTGATCAGGGTCTGC 110

RESULT 15  
 AAX27195/c  
 ID AAX27195 standard; DNA; 306 BP.  
 XX

AC AAX27195;

DT 27-MAY-1999 (first entry)

DE RpoB gene fragment.

XX RpoB gene; mycobacteria; phylogenetic tree construction;

KW mycobacterial species identification; phylogenetic analysis; ss.

XX Mycobacteria intracellulare.

XX WO9905316-A1.

XX 04-FEB-1999.

XX 28-JUL-1998; 98KR-0000228.

XX 28-JUL-1997; 97KR-0035501.

XX (BION-) BIONEER CORP.

XX Kim B, Kook Y;

XX WP1; 1998-539367/46.

XX New pair of polymerase chain reaction (PCR) primers - for  
 PT sequence-specific amplification of the rpoB gene from mycobacterial  
 PT species, useful for detecting and identifying mycobacterial species  
 XX

PS Claim 24; Page 69; 91pp; English.

XX This sequence represents a mycobacterial rpoB gene fragment, that is  
 CC amplified using the PCR primers of the invention. The primers form a  
 CC method of detecting and identifying mycobacterial species by constructing  
 CC a phylogenetic tree for the species. The use of the primers for  
 CC sequence-specific amplification of the rpoB gene (encoding the beta  
 CC subunit of RNA polymerase) from mycobacterial species provides an  
 CC efficient way of characterising these species. In addition to  
 CC phylogenetic analysis, the rpoB gene can be used as an alternative to  
 CC the 16S rRNA gene because it has four subunits, which are highly  
 CC conserved throughout prokaryotes. The method is particularly useful for  
 CC slow growing, fastidious or uncultivable mycobacteria. Also, rifampin  
 CC susceptibility can be simultaneously determined in M. tuberculosis.  
 XX

SQ Sequence 306 BP; 51 A; 111 C; 98 G; 46 T; 0 other;  
 Query Match 100.0%; Score 20; DB 19; Length 306;  
 Best Local Similarity 100.0%; Pred. No. 3.2;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATGTTGATCAGGGTCTGC 20  
 |||||  
 Db 129 GGATGTTGATCAGGGTCTGC 110

Search completed: November 12, 2002, 16:50:34  
 Job time : 14.0983 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 12, 2002, 16:50:46 ; Search time 61.1561 Seconds  
(without alignments)  
9517.553 Million cell updates/sec

Title: US-09-697-123B-26

Perfect score: 20

Sequence: 1 ggaatgtgatcagggtctgc 20

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

1: gb.ba.\*

2: gb.htg.\*

3: gb.in.\*

4: gb.om.\*

5: gb.ov.\*

6: gb.pat.\*

7: gb.ph.\*

8: gb.pl.\*

9: gb.pr.\*

10: gb.ro.\*

11: gb.sts.\*

12: gb.sy.\*

13: gb.un.\*

14: gb.vi.\*

15: em.ba.\*

16: em.fun.\*

17: em.hum.\*

18: em.in.\*

19: em.mu.\*

20: em.om.\*

21: em.or.\*

22: em.ov.\*

23: em.pat.\*

24: em.ph.\*

25: em.pl.\*

26: em.ro.\*

27: em.sts.\*

28: em.un.\*

29: em.vi.\*

30: em.htg\_hum.\*

31: em.htg\_inv.\*

32: em.htg\_other.\*

33: em.htg\_mus.\*

34: em.htg\_pln.\*

35: em.htg\_rod.\*

36: em.htg\_mam.\*

37: em.htg\_vrt.\*

38: em.sy.\*

39: em.htgo\_hum.\*

40: em.htgo\_mus.\*

41: em.htgo\_other.\*

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
C 1	20	100.0	226	6	A47855	A47855 Sequence 69
C 2	20	100.0	254	6	A47823	A47823 Sequence 37
C 3	20	100.0	306	1	AF057449	AF057449 Mycobacte
C 4	20	100.0	306	1	AF057455	AF057455 Mycobacte
C 5	20	100.0	306	1	AF057456	AF057456 Mycobacte
C 6	20	100.0	306	1	AF057457	AF057457 Mycobacte
C 7	20	100.0	306	1	AF057458	AF057458 Mycobacte
C 8	20	100.0	306	1	AF057459	AF057459 Mycobacte
C 9	20	100.0	306	1	AF057460	AF057460 Mycobacte
C 10	20	100.0	306	1	AF057461	AF057461 Mycobacte
C 11	20	100.0	306	1	AF057462	AF057462 Mycobacte
C 12	20	100.0	306	1	AF057463	AF057463 Mycobacte
C 13	20	100.0	306	1	AF057464	AF057464 Mycobacte
C 14	20	100.0	306	1	AF057465	AF057465 Mycobacte
C 15	20	100.0	306	1	AF057467	AF057467 Mycobacte
C 16	20	100.0	306	1	AF057468	AF057468 Mycobacte
C 17	20	100.0	306	1	AF057472	AF057472 Mycobacte
C 18	20	100.0	306	1	AF057477	AF057477 Mycobacte
C 19	20	100.0	306	1	AF057478	AF057478 Mycobacte
C 20	20	100.0	306	1	AF057479	AF057479 Mycobacte
C 21	20	100.0	306	1	AF057480	AF057480 Mycobacte
C 22	20	100.0	306	1	AF057481	AF057481 Mycobacte
C 23	20	100.0	306	1	AF057482	AF057482 Mycobacte
C 24	20	100.0	306	1	AF057483	AF057483 Mycobacte
C 25	20	100.0	306	1	AF057484	AF057484 Mycobacte
C 26	20	100.0	306	1	AF057485	AF057485 Mycobacte
C 27	20	100.0	306	1	AF057487	AF057487 Mycobacte
C 28	20	100.0	306	1	AF057488	AF057488 Mycobacte
C 29	20	100.0	306	1	AF057489	AF057489 Mycobacte
C 30	20	100.0	306	1	AF057490	AF057490 Mycobacte
C 31	20	100.0	306	1	AF057492	AF057492 Mycobacte
C 32	20	100.0	306	1	AF057494	AF057494 Mycobacte
C 33	20	100.0	306	1	AF057495	AF057495 Nocardia
C 34	20	100.0	306	1	AF173087	AF173087 Mycobacte
C 35	20	100.0	306	6	AR157002	AR157002 Sequence
C 36	20	100.0	306	6	AR157004	AR157004 Sequence
C 37	20	100.0	306	6	AR157005	AR157005 Sequence
C 38	20	100.0	306	6	AR157006	AR157006 Sequence
C 39	20	100.0	306	6	AR157009	AR157009 Sequence
C 40	20	100.0	306	6	AR157010	AR157010 Sequence
C 41	20	100.0	306	6	AR157011	AR157011 Sequence
C 42	20	100.0	306	6	AR157012	AR157012 Sequence
C 43	20	100.0	306	6	AR157013	AR157013 Sequence
C 44	20	100.0	306	6	AR157014	AR157014 Sequence
C 45	20	100.0	306	6	AR157015	AR157015 Sequence

#### ALIGNMENTS

RESULT 1  
A47855/c  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL

A47855  
Sequence 69 from Patent WO9533851.  
A47855  
A47855.1 GI:2301741  
unidentified.  
unclassified.  
1 (bases 1 to 226)  
De.B.H., Portael,F., Machtelinckx,L., Jannes,G. and Rossau,R.  
METHOD FOR THE DETECTION OF THE ANTIBIOTIC RESISTANCE SPECTRUM OF  
MYCOBACTERIUM SPECIES  
Patent: WO 9533851-A 69 14-DEC-1995;

226 bp  
DNA  
linear  
PAT 07-MAR-1997

Pred. No. is the number of results predicted by chance to have a



```

/transl_table=11
/product="RNA polymerase beta"
/protein_id="AAD5519.1"
/db_xref="GI:5902498"
/translation="RTVGELIQNIQIRVGMSEVRVVRMTTQDVEAITPOTLINIRP
VVAAIKEFFGTQSOLFMDQNNPLSLTHKRRLSALGPGLSRAGLEVRDVPESH"
BASE COUNT      51 a 106 c 98 g 51 t
ORIGIN

Query Match      100.0%; Score 20; DB 1; Length 306;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGATGTTGATCAGGCTCTGC 20
|||||
Db 129 GGATGTTGATCAGGCTCTGC 110

RESULT 5
AF057456/c
LOCUS      306 bp DNA linear BCT 19-JUN-2000
DEFINITION Mycobacterium aurum RNA polymerase beta (rpoB) gene, partial cds.
ACCESSION  AF057456
VERSION     AF057456.2 GI:8572548
KEYWORDS   .
SOURCE     Mycobacterium aurum.
ORGANISM   Mycobacterium aurum
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Corynebacterineae; Mycobacteriaceae;
Mycobacterium.
REFERENCE  1 (bases 1 to 306)
AUTHORS   Kim,B.J., Lee,S.H., Lyu,M.A., Kim,S.J., Bai,G.H., Chae,G.T.,
Kim,E.C., Cha,C.Y. and Kook,Y.H.
TITLE     Identification of mycobacterial species by comparative sequence
analysis of the RNA polymerase gene (rpoB)
J. Clin. Microbiol. 37 (6), 1714-1720 (1999)
99262756
PUBMED    10325313
REFERENCE  2 (bases 1 to 306)
AUTHORS   Kook,Y.H., Kim,B.J., Lee,S.H., Lyu,M.A., Kim,S.J., Bai,G.H.,
Kim,S.J., Chae,G.T., Kim,E.J. and Cha,C.Y.
TITLE     Direct Submission
JOURNAL   Submitted (06-JUN-2000) Microbiology, Seoul National University
College of Medicine, 28 Yongsong-dong, Chongno-gu, Seoul 110-799,
Korea
REFERENCE  3 (bases 1 to 306)
AUTHORS   Kook,Y.H., Kim,B.J., Lee,S.H., Lyu,M.A., Kim,S.J., Bai,G.H.,
Kim,S.J., Chae,G.T., Kim,E.J. and Cha,C.Y.
TITLE     Direct Submission
JOURNAL   Submitted (19-JUN-2000) Microbiology, Seoul National University
College of Medicine, 28 Yongsong-dong, Chongno-gu, Seoul 110-799,
Korea
REMARK     Sequence update by submitter
COMMENT    On Jun 19, 2000 this sequence version replaced gi:5902499.
FEATURES   Location/Qualifiers
            source
            1..306
            /organism="Mycobacterium aurum"
            /strain="ATCC23366"
            /db_xref="ATCC:23366"
            /db_xref="taxon:1791"
            <1..>306
            /gene="rpoB"
            <1..>306
            /gene="rpoB"
            /codon_start=3
            /transl_table=11
            /product="RNA polymerase beta"
            /protein_id="AAD5520.2"
            /db_xref="GI:8572549"
            /translation="RTVGELIQNIQIRVGLSRMVRVVRMTTQDVEAITPOTLINIRP
            VVAAIKEFFGTQSOLFMDQNNPLSLTHKRRLSALGPGLSRAGLEVRDVPESH"
BASE COUNT      49 a 107 c 98 g 52 t
ORIGIN

```

```

Query Match      100.0%; Score 20; DB 1; Length 306;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGATGTTGATCAGGCTCTGC 20
|||||
Db 129 GGATGTTGATCAGGCTCTGC 110

RESULT 6
AF057457/c
LOCUS      306 bp DNA linear BCT 17-SEP-1999
DEFINITION Mycobacterium avium RNA polymerase beta (rpoB) gene, partial cds.
ACCESSION  AF057457
VERSION     AF057457.1 GI:5902501
KEYWORDS   .
SOURCE     Mycobacterium avium.
ORGANISM   Mycobacterium avium
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Corynebacterineae; Mycobacteriaceae;
Mycobacterium.
REFERENCE  1 (bases 1 to 306)
AUTHORS   Kim,B.J., Lee,S.H., Lyu,M.A., Kim,S.J., Bai,G.H., Chae,G.T.,
Kim,E.C., Cha,C.Y. and Kook,Y.H.
TITLE     Identification of mycobacterial species by comparative sequence
analysis of the RNA polymerase gene (rpoB)
J. Clin. Microbiol. 37 (6), 1714-1720 (1999)
99262756
PUBMED    10325313
REFERENCE  2 (bases 1 to 306)
AUTHORS   Kook,Y.H., Kim,B.J., Lee,S.H., Lyu,M.A., Kim,S.J., Bai,G.H.,
Kim,S.J., Chae,G.T., Kim,E.J. and Cha,C.Y.
TITLE     Direct Submission
JOURNAL   Submitted (06-APR-1998) Microbiology, Seoul National University
College of Medicine, 28 Yongsong-dong, Chongno-gu, Seoul 110-799,
Korea
FEATURES   Location/Qualifiers
            source
            1..306
            /organism="Mycobacterium avium"
            /strain="ATCC25291"
            /db_xref="ATCC:25291"
            /db_xref="taxon:1764"
            <1..>306
            /gene="rpoB"
            <1..>306
            /gene="rpoB"
            /codon_start=3
            /transl_table=11
            /product="RNA polymerase beta"
            /protein_id="AAD5521.1"
            /db_xref="GI:5902502"
            /translation="RTVGELIQNIQIRVGMSEVRVVRMTTQDVEAITPOTLINIRP
            VVAAIKEFFGTQSOLFMDQNNPLSLTHKRRLSALGPGLSRAGLEVRDVPESH"
BASE COUNT      50 a 109 c 101 g 46 t
ORIGIN

Query Match      100.0%; Score 20; DB 1; Length 306;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGATGTTGATCAGGCTCTGC 20
|||||
Db 129 GGATGTTGATCAGGCTCTGC 110

RESULT 7
AF057458/c
LOCUS      306 bp DNA linear BCT 17-SEP-1999
DEFINITION Mycobacterium celatum strain ATCC51131 RNA polymerase beta (rpoB)
gene, partial cds.
ACCESSION  AF057458
VERSION     AF057458.1 GI:5902503

```

KEYWORDS  
SOURCE  
ORGANISM  
Mycobacterium celatum.  
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
Actinomycetales; Corynebacterineae; Mycobacteriaceae;  
Mycobacterium.  
1 (bases 1 to 306)  
Kim,B.J., Lee,S.H., Lyu,M.A., Kim,S.J., Bai,G.H., Chae,G.T.,  
Kim,E.C., Cha,C.Y. and Kook,Y.H.  
Identification of mycobacterial species by comparative sequence  
analysis of the rna polymerase gene (rpoB)  
J. Clin. Microbiol. 37 (6), 1714-1720 (1999)  
JOURNAL  
MEDLINE  
99262756  
PUBMED  
10325313  
REFERENCE  
AUTHORS  
2 (bases 1 to 306)  
Kook,Y.H., Kim,B.J., Lee,S.H., Lyu,M.A., Kim,S.J., Bai,G.H.,  
Kim,S.J., Chae,G.T., Kim,E.J. and Cha,C.Y.  
Direct Submission  
TITLE  
JOURNAL  
College of Medicine, 28 Youngon-dong, Chongno-gu, Seoul 110-799,  
Korea  
FEATURES  
source  
Location/Qualifiers  
1..306  
/organism="Mycobacterium celatum"  
/strain="ATCC51131"  
/db\_xref="ATCC:51131"  
/db\_xref="taxon:28045"  
<1..>306  
/gene="rpoB"  
<1..>306  
/gene="rpoB"  
/codon\_start=3  
/transl\_table=11  
/product="RNA polymerase beta"  
/protein\_id="AAD5522.1"  
/db\_xref="GI:5902504"  
/translation="RTVGELIQNIQVMSMERVVERMTTODVEAITPOTLINIRP  
VVAIKFEFGTSQISQFMDQNNPLSGLTKRRRLNALGPGLSRERAGLEVRDVHPSH"  
BASE COUNT 55 a 102 c 101 g 48 t  
ORIGIN  
Query Match 100.0%; Score 20; DB 1; Length 306;  
Best Local Similarity 100.0%; Pred. No. 26;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 GGATGTTGATCAGGTCCTGC 20  
|||||  
Db 129 GGATGTTGATCAGGTCCTGC 110  
RESULT 8  
AF057459/c  
LOCUS  
DEFINITION  
Mycobacterium celatum strain ATCC51130 RNA polymerase beta (rpoB)  
gene, partial cds.  
ACCESSION  
AF057459  
VERSION  
AF057459.1 GI:5902505  
KEYWORDS  
SOURCE  
ORGANISM  
Mycobacterium celatum.  
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
Actinomycetales; Corynebacterineae; Mycobacteriaceae;  
Mycobacterium.  
1 (bases 1 to 306)  
Kim,B.J., Lee,S.H., Lyu,M.A., Kim,S.J., Bai,G.H., Chae,G.T.,  
Kim,E.C., Cha,C.Y. and Kook,Y.H.  
Identification of mycobacterial species by comparative sequence  
analysis of the rna polymerase gene (rpoB)  
J. Clin. Microbiol. 37 (6), 1714-1720 (1999)  
JOURNAL  
MEDLINE  
99262756  
PUBMED  
10325313  
REFERENCE  
AUTHORS  
2 (bases 1 to 306)  
Kook,Y.H., Kim,B.J., Lee,S.H., Lyu,M.A., Kim,S.J., Bai,G.H.,

Kim,S.J., Chae,G.T., Kim,E.J. and Cha,C.Y.  
Direct Submission  
TITLE  
JOURNAL  
Submitted (06-APR-1998) Microbiology, Seoul National University  
College of Medicine, 28 Youngon-dong, Chongno-gu, Seoul 110-799,  
Korea  
FEATURES  
source  
Location/Qualifiers  
1..306  
/organism="Mycobacterium celatum"  
/strain="ATCC51130"  
/db\_xref="ATCC:51130"  
/db\_xref="taxon:28045"  
<1..>306  
/gene="rpoB"  
<1..>306  
/gene="rpoB"  
/codon\_start=2  
/transl\_table=11  
/product="RNA polymerase beta"  
/protein\_id="AAD5523.1"  
/db\_xref="GI:5902506"  
/translation="SYRRRADPEPDGGRYVEDGAGGPRADDPGRGRRGDHAADPPDOHPS  
GRGGHQGVLRHQPALAVHGPQPIVRADPQASPERARPCWSVPGARPGPRRAPEPL  
"BASE COUNT 55 a 102 c 101 g 48 t  
ORIGIN  
Query Match 100.0%; Score 20; DB 1; Length 306;  
Best Local Similarity 100.0%; Pred. No. 26;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 GGATGTTGATCAGGTCCTGC 20  
|||||  
Db 129 GGATGTTGATCAGGTCCTGC 110  
RESULT 9  
AF057460/c  
LOCUS  
DEFINITION  
Mycobacterium chelonae RNA polymerase beta (rpoB) gene, partial  
cds.  
ACCESSION  
AF057460  
VERSION  
AF057460.1 GI:5902507  
KEYWORDS  
SOURCE  
ORGANISM  
Mycobacterium chelonae.  
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
Actinomycetales; Corynebacterineae; Mycobacteriaceae;  
Mycobacterium.  
1 (bases 1 to 306)  
Kim,B.J., Lee,S.H., Lyu,M.A., Kim,S.J., Bai,G.H., Chae,G.T.,  
Kim,E.C., Cha,C.Y. and Kook,Y.H.  
Identification of mycobacterial species by comparative sequence  
analysis of the rna polymerase gene (rpoB)  
J. Clin. Microbiol. 37 (6), 1714-1720 (1999)  
JOURNAL  
MEDLINE  
99262756  
PUBMED  
10325313  
REFERENCE  
AUTHORS  
2 (bases 1 to 306)  
Kook,Y.H., Kim,B.J., Lee,S.H., Lyu,M.A., Kim,S.J., Bai,G.H.,  
Kim,S.J., Chae,G.T., Kim,E.J. and Cha,C.Y.  
Direct Submission  
TITLE  
JOURNAL  
Submitted (06-APR-1998) Microbiology, Seoul National University  
College of Medicine, 28 Youngon-dong, Chongno-gu, Seoul 110-799,  
Korea  
FEATURES  
source  
Location/Qualifiers  
1..306  
/organism="Mycobacterium chelonae"  
/strain="ATCC35749"  
/db\_xref="ATCC:35749"  
/db\_xref="taxon:1774"  
<1..>306  
/gene="rpoB"  
<1..>306  
/gene="rpoB"

```
/codon_start=3
/transl_table=11
/product="RNA polymerase beta"
/protein_id="AAD5524.1"
/db_xref="GI:5902508"
/translation="RTVGELIQNIIRVGLSRMVRVRRMTTQDVEAITPOTLINIRP
VVAAIKEFFGTQSLSQFMDQNNPLSGLTHKRRLSALGPGGLTRDRAGLEVRDVPSPH"
BASE COUNT      51 a   105 c   94 g   56 t
ORIGIN

Query Match      100.0%; Score 20; DB 1; Length 306;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGATGTTGATCAGGGTCTGC 20
|||||
Db 129 GGATGTTGATCAGGGTCTGC 110

RESULT 10
AF057461/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
1 (bases 1 to 306)
Kim,B.J., Lee,S.H., Lyu,M.A., Kim,S.J., Bai,G.H., Chae,G.T.,
Kim,E.C., Cha,C.Y. and Kook,Y.H.
Identification of mycobacterial species by comparative sequence
analysis of the RNA polymerase gene (rpoB)
J. Clin. Microbiol. 37 (6), 1714-1720 (1999)
99262756
PUBMED
10325313
AUTHORS
Kook,Y.H., Kim,B.J., Lee,S.H., Lyu,M.A., Kim,S.J., Bai,G.H.,
Kim,S.J., Chae,G.T., Kim,E.J. and Cha,C.Y.
Direct Submission
Submitted (06-APR-1998) Microbiology, Seoul National University
College of Medicine, 28 Yongsong-dong, Chongno-gu, Seoul 110-799,
Korea
FEATURES
source
Location/Qualifiers
1..306
/organism="Mycobacterium chitae"
/strain="ATCC19627"
/db_xref="ATCC:19627"
/db_xref="taxon:1792"
<1..>306
/gene="rpoB"
<1..>306
/gene="rpoB"
/codon_start=3
/transl_table=11
/product="RNA polymerase beta"
/protein_id="AAD5525.1"
/db_xref="GI:5902510"
/translation="RTVGELIQNIIRVGLSRMVRVRRMTTQDVEAITPOTLINIRP
VVAAIKEFFGTQSLSQFMDQNNPLSGLTHKRRLSALGPGGLSRERAGLEVRDVPSPH"
BASE COUNT      49 a   108 c   98 g   51 t
ORIGIN

Query Match      100.0%; Score 20; DB 1; Length 306;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGATGTTGATCAGGGTCTGC 20
|||||
Db 129 GGATGTTGATCAGGGTCTGC 110
```

```
RESULT 11
AF057462/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
1 (bases 1 to 306)
Kim,B.J., Lee,S.H., Lyu,M.A., Kim,S.J., Bai,G.H., Chae,G.T.,
Kim,E.C., Cha,C.Y. and Kook,Y.H.
Identification of mycobacterial species by comparative sequence
analysis of the RNA polymerase gene (rpoB)
J. Clin. Microbiol. 37 (6), 1714-1720 (1999)
99262756
PUBMED
10325313
AUTHORS
Kook,Y.H., Kim,B.J., Lee,S.H., Lyu,M.A., Kim,S.J., Bai,G.H.,
Kim,S.J., Chae,G.T., Kim,E.J. and Cha,C.Y.
Direct Submission
Submitted (06-APR-1998) Microbiology, Seoul National University
College of Medicine, 28 Yongsong-dong, Chongno-gu, Seoul 110-799,
Korea
FEATURES
source
Location/Qualifiers
1..306
/organism="Mycobacterium fallax"
/strain="ATCC35219"
/db_xref="ATCC:35219"
/db_xref="taxon:1793"
<1..>306
/gene="rpoB"
<1..>306
/gene="rpoB"
/codon_start=3
/transl_table=11
/product="RNA polymerase beta"
/protein_id="AAD5526.1"
/db_xref="GI:5902512"
/translation="RTVGELIQNIIRVGLSRMVRVRRMTTQDVEAITPOTLINIRP
VVAAIKEFFGTQSLSQFMDQNNPLSGLTHKRRLSALGPGGLSRERAGLEVRDVPSPH"
BASE COUNT      50 a   110 c   101 g   45 t
ORIGIN

Query Match      100.0%; Score 20; DB 1; Length 306;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGATGTTGATCAGGGTCTGC 20
|||||
Db 129 GGATGTTGATCAGGGTCTGC 110

RESULT 12
AF057463/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
1 (bases 1 to 306)
Kim,B.J., Lee,S.H., Lyu,M.A., Kim,S.J., Bai,G.H., Chae,G.T.,
```

```

Kim,E.C., Cha,C.Y. and Kook,Y.H.
Identification of mycobacterial species by comparative sequence
analysis of the RNA polymerase gene (rpoB)
J. Clin. Microbiol. 37 (6), 1714-1720 (1999)
99262756
PUBMED
10325313
REFERENCE
2 (bases 1 to 306)
Kook,Y.H., Kim,B.J., Lee,S.H., Lyu,M.A., Kim,S.J., Bai,G.H.,
Kim,S.J., Chae,G.T., Kim,E.J. and Cha,C.Y.
Direct Submission
Submitted (06-APR-1998) Microbiology, Seoul National University
College of Medicine, 28 Yungno-dong, Chongno-gu, Seoul 110-799,
Korea
FEATURES
    source      Location/Qualifiers
                1..306
                /organism="Mycobacterium flavescens"
                /strain="ATCC14474"
                /db_xref="ATCC:14474"
                /db_xref="taxon:1776"
                <1..>306
                /gene="rpoB"
                <1..>306
                /gene="rpoB"
                /codon_start=3
                /transl_table=11
                /product="RNA polymerase beta"
                /protein_id="AAD5527.1"
                /db_xref="GI:5902514"
                /translation="RTVGELIQNIRVGLSRMRVVRMTTQDVEAITPQTLINEIRP
                VVAAIKEFFGTSQSQFMDQNNPLSLGTHRRLSALPGGLSRERAGLEVRDVHPSH"
BASE COUNT    49 a 106 c 100 g 51 t
ORIGIN
    Query Match      100.0%; Score 20; DB 1; Length 306;
    Best Local Similarity 100.0%; Pred. No. 26;
    Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATGTTGATCAGGTCCTGC 20
    ||||||||||||||||||
Db 129 GGATGTTGATCAGGTCCTGC 110

RESULT 13
AF057464/c
LOCUS
DEFINITION      306 bp DNA linear BCT 17-SEP-1999
Mycobacterium fortuitum strain ATCC6841 RNA polymerase beta (rpoB)
gene, partial cds.
ACCESSION      AF057464
VERSION
KEYWORDS
SOURCE
ORGANISM
Mycobacterium fortuitum.
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Corynebacterineae; Mycobacteriaceae;
Mycobacterium.
REFERENCE
1 (bases 1 to 306)
Kim,B.J., Lee,S.H., Lyu,M.A., Kim,S.J., Bai,G.H., Chae,G.T.,
Kim,E.C., Cha,C.Y. and Kook,Y.H.
Identification of mycobacterial species by comparative sequence
analysis of the RNA polymerase gene (rpoB)
J. Clin. Microbiol. 37 (6), 1714-1720 (1999)
99262756
PUBMED
10325313
REFERENCE
2 (bases 1 to 306)
Kook,Y.H., Kim,B.J., Lee,S.H., Lyu,M.A., Kim,S.J., Bai,G.H.,
Kim,S.J., Chae,G.T., Kim,E.J. and Cha,C.Y.
Direct Submission
Submitted (06-APR-1998) Microbiology, Seoul National University
College of Medicine, 28 Yungno-dong, Chongno-gu, Seoul 110-799,
Korea
FEATURES
    source      Location/Qualifiers
                1..306
                /organism="Mycobacterium fortuitum"
                /strain="ATCC6841"
                /db_xref="ATCC:6841"
                /db_xref="taxon:1766"
                <1..>306
                /gene="rpoB"
                <1..>306
                /gene="rpoB"
                /codon_start=3
                /transl_table=11
                /product="RNA polymerase beta"
                /protein_id="AAD5529.1"
                /db_xref="GI:5902518"
                /translation="RTVGELIQNIRVGLSRMRVVRMTTQDVEAITPQTLINEIRP
                VVAAIKEFFGTSQSQFMDQNNPLSLGTHRRLSALPGGLSRERAGLEVRDVHSH"
BASE COUNT    49 a 106 c 98 g 53 t
ORIGIN
    Query Match      100.0%; Score 20; DB 1; Length 306;
    Best Local Similarity 100.0%; Pred. No. 26;
    Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATGTTGATCAGGTCCTGC 20
    ||||||||||||||||||
Db 129 GGATGTTGATCAGGTCCTGC 110

RESULT 14
AF057465/c
LOCUS
DEFINITION      306 bp DNA linear BCT 17-SEP-1999
Mycobacterium fortuitum strain ATCC49403 RNA polymerase beta (rpoB)
gene, partial cds.
ACCESSION      AF057465
VERSION
KEYWORDS
SOURCE
ORGANISM
Mycobacterium fortuitum.
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Corynebacterineae; Mycobacteriaceae;
Mycobacterium.
REFERENCE
1 (bases 1 to 306)
Kim,B.J., Lee,S.H., Lyu,M.A., Kim,S.J., Bai,G.H., Chae,G.T.,
Kim,E.C., Cha,C.Y. and Kook,Y.H.
Identification of mycobacterial species by comparative sequence
analysis of the RNA polymerase gene (rpoB)
J. Clin. Microbiol. 37 (6), 1714-1720 (1999)
99262756
PUBMED
10325313
REFERENCE
2 (bases 1 to 306)
Kook,Y.H., Kim,B.J., Lee,S.H., Lyu,M.A., Kim,S.J., Bai,G.H.,
Kim,S.J., Chae,G.T., Kim,E.J. and Cha,C.Y.
Direct Submission
Submitted (06-APR-1998) Microbiology, Seoul National University
College of Medicine, 28 Yungno-dong, Chongno-gu, Seoul 110-799,
Korea
FEATURES
    source      Location/Qualifiers
                1..306
                /organism="Mycobacterium fortuitum"
                /strain="ATCC49403"
                /db_xref="ATCC:49403"
                /db_xref="taxon:1766"
                <1..>306
                /gene="rpoB"
                <1..>306
                /gene="rpoB"
                /codon_start=3
                /transl_table=11
                /product="RNA polymerase beta"
                /protein_id="AAD5529.1"
                /db_xref="GI:5902518"
                /translation="RTVGELIQNIRVGLSRMRVVRMTTQDVEAITPQTLINEIRP
                VVAAIKEFFGTSQSQFMDQNNPLSLGTHRRLSALPGGLSRERAGLEVRDVHSH"
BASE COUNT    49 a 106 c 98 g 53 t
ORIGIN

```

```

Kim,E.C., Cha,C.Y. and Kook,Y.H.
Identification of mycobacterial species by comparative sequence
analysis of the RNA polymerase gene (rpoB)
J. Clin. Microbiol. 37 (6), 1714-1720 (1999)
99262756
PUBMED
10325313
REFERENCE
2 (bases 1 to 306)
Kook,Y.H., Kim,B.J., Lee,S.H., Lyu,M.A., Kim,S.J., Bai,G.H.,
Kim,S.J., Chae,G.T., Kim,E.J. and Cha,C.Y.
Direct Submission
Submitted (06-APR-1998) Microbiology, Seoul National University
College of Medicine, 28 Yungno-dong, Chongno-gu, Seoul 110-799,
Korea
FEATURES
    source      Location/Qualifiers
                1..306
                /organism="Mycobacterium flavescens"
                /strain="ATCC14474"
                /db_xref="ATCC:14474"
                /db_xref="taxon:1776"
                <1..>306
                /gene="rpoB"
                <1..>306
                /gene="rpoB"
                /codon_start=3
                /transl_table=11
                /product="RNA polymerase beta"
                /protein_id="AAD5527.1"
                /db_xref="GI:5902514"
                /translation="RTVGELIQNIRVGLSRMRVVRMTTQDVEAITPQTLINEIRP
                VVAAIKEFFGTSQSQFMDQNNPLSLGTHRRLSALPGGLSRERAGLEVRDVHPSH"
BASE COUNT    49 a 106 c 100 g 51 t
ORIGIN
    Query Match      100.0%; Score 20; DB 1; Length 306;
    Best Local Similarity 100.0%; Pred. No. 26;
    Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATGTTGATCAGGTCCTGC 20
    ||||||||||||||||||
Db 129 GGATGTTGATCAGGTCCTGC 110

RESULT 13
AF057464/c
LOCUS
DEFINITION      306 bp DNA linear BCT 17-SEP-1999
Mycobacterium fortuitum strain ATCC6841 RNA polymerase beta (rpoB)
gene, partial cds.
ACCESSION      AF057464
VERSION
KEYWORDS
SOURCE
ORGANISM
Mycobacterium fortuitum.
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Corynebacterineae; Mycobacteriaceae;
Mycobacterium.
REFERENCE
1 (bases 1 to 306)
Kim,B.J., Lee,S.H., Lyu,M.A., Kim,S.J., Bai,G.H., Chae,G.T.,
Kim,E.C., Cha,C.Y. and Kook,Y.H.
Identification of mycobacterial species by comparative sequence
analysis of the RNA polymerase gene (rpoB)
J. Clin. Microbiol. 37 (6), 1714-1720 (1999)
99262756
PUBMED
10325313
REFERENCE
2 (bases 1 to 306)
Kook,Y.H., Kim,B.J., Lee,S.H., Lyu,M.A., Kim,S.J., Bai,G.H.,
Kim,S.J., Chae,G.T., Kim,E.J. and Cha,C.Y.
Direct Submission
Submitted (06-APR-1998) Microbiology, Seoul National University
College of Medicine, 28 Yungno-dong, Chongno-gu, Seoul 110-799,
Korea
FEATURES
    source      Location/Qualifiers
                1..306
                /organism="Mycobacterium fortuitum"
                /strain="ATCC6841"
                /db_xref="ATCC:6841"
                /db_xref="taxon:1766"
                <1..>306
                /gene="rpoB"
                <1..>306
                /gene="rpoB"
                /codon_start=3
                /transl_table=11
                /product="RNA polymerase beta"
                /protein_id="AAD5529.1"
                /db_xref="GI:5902518"
                /translation="RTVGELIQNIRVGLSRMRVVRMTTQDVEAITPQTLINEIRP
                VVAAIKEFFGTSQSQFMDQNNPLSLGTHRRLSALPGGLSRERAGLEVRDVHSH"
BASE COUNT    49 a 106 c 98 g 53 t
ORIGIN

```



Query Match 100.0%; Score 20; DB 1; Length 306;  
 Best Local Similarity 100.0%; Pred. No. 26;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATGTTGATCAGGCTCTGC 20  
 |||  
 Db 129 GGATGTTGATCAGGCTCTGC 110

## RESULT 15

AF057467/c  
 LOCUS  
 DEFINITION Mycobacterium genavense RNA polymerase beta (rpoB) gene, partial cds.  
 ACCESSION AF057467  
 VERSION AF057467.1 GI:5902521  
 KEYWORDS  
 SOURCE  
 ORGANISM Mycobacterium genavense.  
 Mycobacterium genavense  
 Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 Actinomycetales; Corynebacterineae; Mycobacteriaceae;  
 Mycobacterium.  
 REFERENCE 1 (bases 1 to 306)  
 Kim,B.J., Lee,S.H., Lyu,M.A., Kim,S.J., Bai,G.H., Chae,G.T.,  
 Kim,E.C., Cha,C.Y. and Kook,Y.H.  
 Identification of mycobacterial species by comparative sequence  
 analysis of the RNA polymerase gene (rpoB)  
 J. Clin. Microbiol. 37 (6), 1714-1720 (1999)  
 99262756  
 MEDLINE  
 PUBMED 10325313  
 REFERENCE 2 (bases 1 to 306)  
 Kook,Y.H., Kim,B.J., Lee,S.H., Lyu,M.A., Kim,S.J., Bai,G.H.,  
 Kim,S.J., Chae,G.T., Kim,E.J. and Cha,C.Y.  
 Direct Submission  
 Submitted (06-APR-1998) Microbiology, Seoul National University  
 College of Medicine, 28 Youngon-dong, Chongno-gu, Seoul 110-799,  
 Korea

FEATURES  
source

Location/Qualifiers  
 1..306  
 /organism="Mycobacterium genavense"  
 /strain="ATCC51233"  
 /db\_xref="ATCC:51233"  
 /db\_xref="taxon:36812"  
 <1..>306  
 /gene="rpoB"  
 <1..>306  
 /gene="rpoB"  
 /codon\_start=3  
 /transl\_table=11  
 /product="RNA polymerase beta"  
 /protein\_id="AAD5531.1"  
 /db\_xref="GI:5902521"  
 /translation="RTVGLIQIRVGMSEVVRMTTQDVEATPOTFLINIRP  
 VVAATKEFGTSOLFMDONNPLSGLTHKRRLSALPGGLSRERAGLEVRDVFPSH#  
 51 a 99 c 105 g 51 t

BASE COUNT  
ORIGIN

Query Match 100.0%; Score 20; DB 1; Length 306;  
 Best Local Similarity 100.0%; Pred. No. 26;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATGTTGATCAGGCTCTGC 20  
 |||  
 Db 129 GGATGTTGATCAGGCTCTGC 110

Search completed: November 13, 2002, 01:28:16  
 Job time : 63.8561 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 12, 2002, 16:59:42 ; Search time 101.315 Seconds  
(without alignments)  
3037.202 Million cell updates/sec

Title: US-09-697-123b-25

Perfect score: 19  
Sequence: 1 tcaaggagaaagcgctacga 19

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

- EST:\*
- 1: em\_estba:\*
  - 2: em\_esthum:\*
  - 3: em\_estlin:\*
  - 4: em\_estmu:\*
  - 5: em\_estcov:\*
  - 6: em\_estpl:\*
  - 7: em\_estro:\*
  - 8: em\_htc:\*
  - 9: gb\_est1:\*
  - 10: gb\_est2:\*
  - 11: gb\_htc:\*
  - 12: gb\_est3:\*
  - 13: gb\_est4:\*
  - 14: gb\_est5:\*
  - 15: em\_estfun:\*
  - 16: em\_estom:\*
  - 17: gb\_gss:\*
  - 18: em\_gss\_hum:\*
  - 19: em\_gss\_inv:\*
  - 20: em\_gss\_pln:\*
  - 21: em\_gss\_vrt:\*
  - 22: em\_gss\_fun:\*
  - 23: em\_gss\_man:\*
  - 24: em\_gss\_mus:\*
  - 25: em\_gss\_other:\*
  - 26: em\_gss\_pro:\*
  - 27: em\_gss\_rod:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
c 1	19	100.0	1282	9	AI770311 42 Mycoba
c 2	18	94.7	705	17	AQ575881 nbxb0088A
c 3	17.4	91.6	609	12	BF495134 AT03975.5
c 4	17.4	91.6	647	13	BM427436 pgfzn.pk0
c 5	17.4	91.6	659	12	BF504233 AT05552.5
c 6	17.4	91.6	663	12	BF498268 AT12651.5

7	17.4	91.6	667	9	AJ451042
c 8	17	89.5	558	17	BH785787
9	16.4	86.3	249	17	BH424101
c 10	16.4	86.3	265	10	BE517162
c 11	16.4	86.3	290	9	AI948950
12	16.4	86.3	375	17	BH484296
13	16.4	86.3	432	14	H06751
14	16.4	86.3	437	13	BM375873
c 15	16.4	86.3	441	17	BH717250
16	16.4	86.3	509	12	BG714670
17	16.4	86.3	567	13	B1168616
18	16.4	86.3	579	13	BM377270
19	16.4	86.3	596	14	BQ465913
20	16.4	86.3	623	13	B1780229
21	16.4	86.3	677	14	BQ466189
22	16.4	86.3	692	9	AL598382
c 23	16.4	86.3	715	17	AG086517
24	16.4	86.3	730	13	B1568344
25	16.4	86.3	747	12	BG720538
c 26	16.4	86.3	806	17	BH428934
27	16.4	86.3	860	17	BH590938
28	16.4	86.3	863	12	BG368660
29	16.4	86.3	863	13	B1547968
30	16.4	86.3	871	9	AL547893
31	16.4	86.3	924	13	B1546826
32	16.4	86.3	947	12	BG365925
33	16.4	86.3	964	14	BQ934678
34	16.4	86.3	999	17	CNS040UL
35	16	84.2	212	14	T23183
36	15.8	83.2	127	10	AW787398
37	15.8	83.2	168	13	BM517883
38	15.8	83.2	172	13	BM516341
39	15.8	83.2	227	10	AW144977
40	15.8	83.2	229	10	AW724941
41	15.8	83.2	229	10	AW724942
42	15.8	83.2	244	14	BQ441968
43	15.8	83.2	267	17	BH231978
44	15.8	83.2	280	10	AW718755
45	15.8	83.2	280	10	AW718756

ALIGNMENTS

RESULT 1  
AI770311/c  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
COMMENT

AI770311 1282 bp mRNA linear EST 24-JAN-2000  
42 Mycobacterium anaerobic stationary phase library Mycobacterium  
smegmatis CDNA, mRNA sequence.  
AI770311 GI:6742680  
EST.  
Mycobacterium smegmatis.  
Mycobacterium smegmatis  
Bacteria: Actinobacteria; Actinobacteridae; Actinomycetales;  
Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
1 (bases 1 to 1282)  
Murugasu-Oel, B., Tay, A. and Dick, T.  
Upregulation of stress response genes and ABC transporters in  
anaerobic stationary-phase Mycobacterium smegmatis  
Mol. Gen. Genet. 262 (4-5), 677-682 (1999)  
20092472  
Contact: Murugasu-Oel, B.  
Mycobacterium Laboratory  
Institute of Molecular and Cell Biology  
30 Medical Drive, Singapore 117609, Republic of Singapore  
Tel: 65 874 3011  
Fax: 65 779 1117  
Email: mcbom@imcb.nus.edu.sg  
Insert Length: 1282 Std Error: 0.00  
Seq primer: T3 Forward; T7 Backward.

FEATURES  
Location/Qualifiers  
1..1282  
source

```

/organism="Mycobacterium smegmatis"
/strain="mc2155"
/db_xref="taxon:1772"
/clone_lib="Mycobacterium anaerobic stationary phase
library"
/lab_host="E. coli XL1-Blue MRF"
/notes="Vector: Lambda ZAP II; Bacilli were disrupted using
a Mini bead beater (Biospec). RNA was isolated using the
RNeasy protocol (Qiagen). Purified RNA was subjected to 2
rounds of digestion with RNase-free DNase I (Promega).
DNase I was heat-inactivated at 75oC for 5 min. and
removed by using RNeasy columns followed by phenol
extraction and ethanol precipitation. The RNA
preparations were confirmed to be free of genomic DNA
contamination by carrying out PCR and RT-PCR using the
Access kit (Promega) and primers specific for the
histone-like protein gene hlp (Lee et al., 1998). cDNA
was synthesized using random hexamer primers (Promega)
and Stratagene's cDNA synthesis kit. cDNA fragments were
ligated into lambda ZAP II vector and packaged in vitro
using Stratagene reagents."
BASE COUNT      230 a 395 c 405 g 252 t
ORIGIN
Query Match      100.0%; Score 19; DB 9; Length 1282;
Best Local Similarity 100.0%; Pred. No. 99;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCAAGGAGAGCGCTACGA 19
|||||
Db 236 TCAAGGAGAGCGCTACGA 218

RESULT 2
AQ575881/c
LOCUS      705 bp DNA linear GSS 02-JUN-1999
DEFINITION nbxb0088A06r CUGI Rice BAC Library Oryza sativa genomic clone
ACCESSION nbxb0088A06r, DNA sequence.
VERSION    AQ575881
KEYWORDS   AQ575881.1 GI:4976366
SOURCE     GSS.
ORGANISM   Oryza sativa.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
1 (bases 1 to 705)
Wing, R.A. and Dean, R.A.
A BAC End Sequencing Framework to Sequence the Rice Genome
Unpublished (1998)
Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
Seq primer: GGAACAGCTATGACCATG
Class: BAC ends
High quality sequence start: 62
High quality sequence stop: 336.
Location/Qualifiers
1. 705
/organism="Oryza sativa"
/strain="Japonica"
/cultivar="Nipponbare"
/db_xref="taxon:4530"
/clone_lib="nbxb0088A06r"
/tissue_type="Leaf"
/lab_host="E. coli DH10B"
/notes="vector: pBeloBAC11; Site_1: HindIII; Site_2:
HindIII; Rice is one of two most popular grains in the

```

```

world. Half of the world population especially those
inhabiting highly populated areas of the humid tropics
and subtropics, rely on rice as their primary source of
carbohydrate. Monocotyledonous rice is a diploid plant
(2n=24) with a haploid genome equivalent of 431 Mbp
(Arumuganathan and Earle, 1991). The relatively small
genome of rice, three times larger than that of
Arabidopsis, makes it suitable for genomic studies. In
order to facilitate positional cloning, physical mapping
and genome sequencing of rice, we have constructed a BAC
library from Oryza sativa, Nipponbare variety. The
library contains 36,864 clones with an average insert size
of 128.5 Kb providing 10.9 haploid genome equivalents. The
deep coverage allows the isolation a particular sequence
with a probability of 99.9 %. Two high density filters,
each containing 18 432 clones (doubly spotted), represent
the whole library for colony screening.
BASE COUNT      190 a 172 c 158 g 181 t 4 others
ORIGIN
Query Match      94.7%; Score 18; DB 17; Length 705;
Best Local Similarity 100.0%; Pred. No. 2,4e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCAAGGAGAGCGCTACG 18
|||||
Db 94 TCAAGGAGAGCGCTACG 77

RESULT 3
BF495134/c
LOCUS      609 bp mRNA linear EST 19-APR-2001
DEFINITION AT03975.5prime AT Drosophila melanogaster adult testes pOTB7
Drosophila melanogaster cDNA clone AT03975 5 similar to Med:
FBan0001775 'transcription factor' located on: 3R 100D1-100D1;:
04/07/2001, mRNA sequence.
ACCESSION BF495134
VERSION    BF495134.2 GI:13685621
KEYWORDS   EST.
SOURCE     fruit fly.
ORGANISM   Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 609)
Stapleton, M., Brokstein, P., Hong, L., Agbayani, A., Baxter, E., Berman
, D., Carlson, J., Champe, M., Chavez, C., Chew, M., Dorsett, V., Farfan
, D., Frise, E., George, R., Gonzalez, M., Guarin, H., Harris, N., Li, P.,
Liao, G., Miranda, A., Misra, S., Mungall, C.J., Nunoo, J., Pacלב, J.,
Paragas, V., Park, S., Phouanavong, S., Wan, K., Yu, C., Lewis, S.E.,
Ceiniker, S. and Rubin, G.M.
BDGP/HHMI AT Drosophila EST Project
Unpublished (2000)
On Dec 6, 2000 this sequence version replaced gi:11578435.
Contact: Stapleton, M.
BDGP
Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
hit genomic AE003779; arm:3R [27296749,27528812]
estimated-cyto:100D1-100E3: 04/07/2001
Plate: AT.39 row: G column: 3
High quality sequence stop: 595.
Location/Qualifiers
1. 609
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone_lib="AT03975"
/clone_lib="AT Drosophila melanogaster adult testes pOTB7"
/dev_stage="0-3 day old Ore-R males"
/lab_host="Plates AT.10-AT.120: DH5-alpha. Plates

```

```

AT.121-AT.319; DH5-alpha Tona"
/note="Organ: ADULT testes; Vector: pOTB7; Site_1: EcoRI;
Site_2: XhoI; The mRNA for the testis library was made
from testes and seminal vesicles hand dissected from 0-3
day old Ore-R males. RNA kindly provided by the lab of
Margaret Fuller. Sized fractionated cDNAs were directly
ligated into pOTB7. Plasmid cDNA library."
BASE COUNT      120 a   167 c   188 g   134 t
ORIGIN

Query Match      91.6%; Score 17.4; DB 12; Length 609;
Best Local Similarity 94.7%; Pred. No. 4.5e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TCAAGGAGACGGCTACGA 19
|||||
Db 159 TCAAGGAGACGGCGACGA 141

RESULT 4
BM427436
LOCUS
DEFINITION
pgf2n.pk006.m12 Normalized Chicken Abdominal Fat Library (pgf2n)
Gallus gallus cDNA clone pgf2n.pk006.m12 5' similar to gi|14091760
ref|NP_114461.1| APP-binding protein 1 [Rattus norvegicus]
gb|AAD09247.1| (U90829) APP-binding protein 1 [Rattus norvegicus],
mRNA sequence.
BM427436
VERSION
BM427436.1 GI:18432610
KEYWORDS
EST.
ORGANISM
Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
REFERENCE
1 (bases 1 to 647)
Cogburn,L.A., Morgan,R. and Burnside,J.
ESTs from Normalized Chicken fat cDNA library-USDA/IFAFS Animal
Genome Project
JOURNAL
Unpublished (2002)
CONTACT: Larry A. Cogburn
University of Delaware
Townsend Hall, Newark, DE 19717, USA
Tel: 302-831-1335
Fax: 302-831-2822
Email: cogburnudel.edu, www.chickest.udel.edu.
FEATURES
source
1..647
/organism="Gallus gallus"
/strain="Commercial broiler, Ottawa Research Centre,
leghorn"
/db_xref="taxon:9031"
/clone="pgf2n.pk006.m12"
/clone_lib="Normalized Chicken Abdominal Fat Library
(pg2n)"
/sex="Male and Female"
/tissue_type="Abdominal Fat"
/dev_stage="Embryonic (d18,d19); post-hatch (d1.w3,w7,w9
,w16.lyr)"
/lab_host="E. coli EMDH10B"
/note="Vector: pCMVSPORT6; Library made from equivalent
pools of total RNA isolated from each developmental age
(across strains); Single pass sequencing from 5'-end"
BASE COUNT      177 a   133 c   159 g   165 t   13 others
ORIGIN

Query Match      91.6%; Score 17.4; DB 13; Length 647;
Best Local Similarity 94.7%; Pred. No. 4.6e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TCAAGGAGACGGCTACGA 19
|||||
Db 40 TCAAGGAGACGGCGTACGA 58

AT.121-AT.319; DH5-alpha Tona"
/note="Organ: ADULT testes; Vector: pOTB7; Site_1: EcoRI;
Site_2: XhoI; The mRNA for the testis library was made
from testes and seminal vesicles hand dissected from 0-3
day old Ore-R males. RNA kindly provided by the lab of
Margaret Fuller. Sized fractionated cDNAs were directly
ligated into pOTB7. Plasmid cDNA library."
BASE COUNT      120 a   167 c   188 g   134 t
ORIGIN

Query Match      91.6%; Score 17.4; DB 12; Length 609;
Best Local Similarity 94.7%; Pred. No. 4.5e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TCAAGGAGACGGCTACGA 19
|||||
Db 159 TCAAGGAGACGGCGACGA 141

RESULT 5
BF504233/c
LOCUS
DEFINITION
BF504233
659 bp mRNA linear EST 19-APR-2001
AT05552.5 Prime AT Drosophila melanogaster adult testes pOTB7
Drosophila melanogaster cDNA clone AT05552 5 similar to Med:
FBan0001775 'transcription factor' located on: 3R 100D1-100D1.;
04/07/2001, mRNA sequence.
BF504233
VERSION
BF504233.2 GI:13686863
KEYWORDS
EST.
SOURCE
fruit fly.
ORGANISM
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
1 (bases 1 to 659)
Stapleton,M., Brokstein,P., Hong,L., Agbayani,A., Baxter,E., Berman
,B., Carlson,J., Champe,M., Chavez,C., Chew,M., Dorsett,V., Farfan
,D., Frise,E., George,R., Gonzalez,M., Guarin,H., Harris,N., Li,P.,
Liao,G., Miranda,A., Misra,S., Mungall,C.J., Nunoo,J., Pacleb,J.,
Paragas,V., Park,S., Phouanavong,S., Wan,K., Yu,C., Lewis,S.E.,
Celnikier,S. and Rubin,G.M.
BDGP/HMI AT Drosophila EST Project
Unpublished (2000)
JOURNAL
On Dec 6, 2000 this sequence version replaced gi:11587534.
CONTACT: Stapleton, M.
BDGP
Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST, estfruitfly.berkeley.edu
hit genomic AE003779: arm:3R [27296749..27528812]
estimated-cyto:100D1-100E3: 04/07/2001
Plate: AT.55 row: E column: 4
High quality sequence stop: 657.
Location/Qualifiers
1..659
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone="AT05552"
/clone_lib="AT Drosophila melanogaster adult testes pOTB7"
/sex="male"
/dev_stage="0-3 day old Ore-R males"
/lab_host="Plates AT.10-AT.120; DH5-alpha. Plates
AT.121-AT.319; DH5-alpha Tona"
/note="Organ: ADULT testes; Vector: pOTB7; Site_1: EcoRI;
Site_2: XhoI; The mRNA for the testis library was made
from testes and seminal vesicles hand dissected from 0-3
day old Ore-R males. RNA kindly provided by the lab of
Margaret Fuller. Sized fractionated cDNAs were directly
ligated into pOTB7. Plasmid cDNA library."
BASE COUNT      128 a   177 c   213 g   141 t
ORIGIN

Query Match      91.6%; Score 17.4; DB 12; Length 659;
Best Local Similarity 94.7%; Pred. No. 4.6e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TCAAGGAGACGGCTACGA 19
|||||
Db 159 TCAAGGAGACGGCGACGA 141

RESULT 6
BF498268/c
LOCUS
DEFINITION
BF498268
663 bp mRNA linear EST 19-APR-2001
AT12651.5 Prime AT Drosophila melanogaster adult testes pOTB7
Drosophila melanogaster cDNA clone AT12651 5 similar to Med:
FBan0001775 'transcription factor' located on: 3R 100D1-100D1.;
04/09/2001, mRNA sequence.
BF498268
ACCESSION

```

```

VERSION      BF498268.2  GI:13690169
KEYWORDS
SOURCE       fruit fly.
ORGANISM     Drosophila melanogaster

REFERENCE    1 (bases 1 to 663)
AUTHORS      Stapleton, M., Brokstein, P., Hong, L., Agbayani, A., Baxter, E., Berman
              , B., Carlson, J., Champe, M., Chavez, C., Chew, M., Dorsett, V., Farfan
              , D., Frise, E., George, R., Gonzalez, M., Guarin, H., Harris, N., Li, P.,
              Liao, G., Miranda, A., Mista, S., Mungall, C.J., Nunoo, J., Pacleb, J.,
              Paragas, V., Park, S., Phouanavong, S., Wan, K., Yu, C., Lewis, S.E.,
              Celniker, S. and Rubin, G.M.
TITLE        BDGP/HMI AT Drosophila EST Project
JOURNAL      Unpublished (2000)
COMMENT      On Dec 6, 2000 this sequence version replaced gi:11581569.
              Contact: Stapleton, M.
              BDGP

              Lawrence Berkeley National Lab
              One Cyclotron Rd, Berkeley, CA 94720, USA
              Fax: 510 486 6798
              Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
              hit genomic AE003779: arm:3R [27296749,27528812]
              estimated-cyto:100D1-100E3: 04/09/2001
              Plate: AT.126 row: E column: 3
              High quality sequence stop: 624.
              Location/Qualifiers
                1..663
                  /organism="Drosophila melanogaster"
                  /db_xref="taxon:7227"
                  /clone="AT12651"
                  /sex="male"
                  /dev_stage="0-3 day old Ore-R males"
                  /lab_host="Plates AT.10-AT.120: DH5-alpha. Plates
                  AT.121-AT.319: DH5-alpha Tona"
                  /note="Organ: ADULT testes; Vector: pOTB7; Site_1: EcoRI;
                  Site_2: XhoI; The mRNA for the testis library was made
                  from testes and seminal vesicles hand dissected from 0-3
                  day old Ore-R males. RNA kindly provided by the lab of
                  Margaret Fuller. Sized fractionated cDNAs were directly
                  ligated into pOTB7. Plasmid cDNA library."
BASE COUNT   128 a 178 c 216 g 141 t
ORIGIN
FEATURES     source
  Query Match      91.6%; Score 17.4; DB 12; Length 663;
  Best Local Similarity 94.7%; Pred. No. 4.6e+02;
  Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TCAAGGAGAGCGCTACGA 19
|||||
Db 159 TCAAGGAGAGCGCGACGA 141

RESULT 7
AJ451042
LOCUS        667 bp mRNA linear EST 19-APR-2002
DEFINITION   riken1 Gallus gallus cDNA clone 2756r1, mRNA sequence.
ACCESSION    AJ451042
VERSION      AJ451042.1 GI:20218263
KEYWORDS     EST.
SOURCE       chicken.
ORGANISM     Gallus gallus
              Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
              Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
              Phasianinae; Gallus.
              1 (bases 1 to 667)
              Buerstedde, J.M.
              Gallus gallus bursal lymphocyte EST
              Unpublished (2002)
              Contact: Buerstedde JM
              Cellular Immunology

HEINRICH-PETTE-INSTITUTE
MARTINSTR. 52, 20251 HAMBURG, GERMANY
EMAIL: URL: http://genetics.hpi.uni-hamburg.de/dt40est.html.

FEATURES     source
  Location/Qualifiers
    /organism="Gallus gallus"
    /db_xref="taxon:9031"
    /clone="2756r1"
    /clone_lib="riken1"
    /cell_type="bursal lymphocyte"
    /dev_stage="2-3 weeks old"
    /note="CB inbred strain"
BASE COUNT   186 a 141 c 170 g 169 t
ORIGIN
  Query Match      91.6%; Score 17.4; DB 9; Length 667;
  Best Local Similarity 94.7%; Pred. No. 4.6e+02;
  Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TCAAGGAGAGCGCTACGA 19
|||||
Db 59 TCAAGGAGAGCGCTACGA 77

RESULT 8
BH785787/c
LOCUS        558 bp DNA linear GSS 28-MAR-2002
DEFINITION   fzm013f050e10k0 fzm0 filtered library Zea mays genomic clone
              BH785787
              fzm013f050e10 5', DNA sequence.
ACCESSION    BH785787
VERSION      BH785787.1 GI:19790828
KEYWORDS     GSS.
SOURCE       Zea mays.
              ORGANISM
              Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
              clade: Panicoideae; Andropogoneae; Zea.
              1 (bases 1 to 558)
              Budiman, M.A., Freese, R.G., Bedell, J.A., Nunberg, A.N. and Lakey, N.D.
              Genefresher methylation filtered genomic sequences from maize
              Unpublished (2002)
              Contact: Bedell JA
              Orion Genomics, LLC
              4041 Forest Park Ave, St. Louis, MO 63108, USA
              Tel: 314 615 6979
              Fax: 314 615 5975
              Email: jbedell@oriongenomics.com
              Plate: fzm013f050 row: e column: 10
              Seq primer: SK reverse
              Class: shotgun
              High quality sequence stop: 558.
              Location/Qualifiers
                1..558
                  /organism="Zea mays"
                  /cultivar="MO17"
                  /db_xref="taxon:4577"
                  /clone="fzm013f050e10"
                  /clone_lib="fzm0 filtered library"
                  /note="Organ: leaf; Vector: pBCSK(-); Site_1: HincII; DNA
                  prepared from purified nuclei was randomly sheared,
                  end-repaired, size fractionated to enrich for the 0.5 to
                  5 kb fraction, ligated into HincII-digested pBCSK(-)
                  vector and electroporated into E. coli cells."
BASE COUNT   134 a 124 c 141 g 159 t
ORIGIN
  Query Match      89.5%; Score 17; DB 17; Length 558;
  Best Local Similarity 100.0%; Pred. No. 6.8e+02;
  Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 AAGGAGAGCGCTACGA 19
|||||
Db 453 AAGGAGAGCGCTACGA 437

```

## RESULT 9

BH424101  
 LOCUS  
 DEFINITION BOGOY49TR BOGO Brassica oleracea genomic clone BOGOY49, DNA linear GSS 12-DEC-2001 sequence.  
 ACCESSION BH424101  
 VERSION BH424101.1 GI:17609829  
 KEYWORDS GSS.  
 SOURCE Brassica oleracea.  
 ORGANISM Brassica oleracea.  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.  
 1 (bases 1 to 249)  
 Town,C.D., Van Aken,S., Utterback,T. and Fraser,C.M.  
 Whole genome shotgun sequencing of Brassica oleracea  
 Unpublished (2001)  
 Other\_GSSs: BOGOY49TF  
 Contact: Chris Town  
 TIGR  
 9712 Medical Center Drive, Rockville, MD 20850, USA.  
 Tel: 301-838-3523  
 Fax: 301-838-0208  
 Email: cdtown@tigr.org  
 DNA is from a doubled haploid provided by Tom Osborn.  
 Seq primer: rR  
 Class: sheared ends.

## FEATURES

source

1..249  
 /organism="Brassica oleracea"  
 /strain="T0100DH3"  
 /db\_xref="taxon:3712"  
 /clone="BOGOY49"  
 /clone\_lib="BOGO"  
 /note="Vector: pHD1; Site\_1: BstXI; 2-3 kb sheared genomic DNA inserted into pHD1 using BstXI linkers"  
 BASE COUNT 77 a 47 c 56 g 59 t  
 ORIGIN

Query Match 86.3%; Score 16.4; DB 17; Length 249;  
 Best Local Similarity 94.4%; Pred. No. 9.8e+02;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 2 CAAGGAGAGCGGTACGA 19  
 ||||| ||||| |||||  
 Db 60 CAAGGAGAGCGGTACGA 77

## RESULT 10

BES17162/c  
 LOCUS  
 DEFINITION WHE0615\_H12\_P232A Wheat ABA-treated embryo cDNA library Triticum aestivum cDNA clone WHE0615\_H12\_P23, mRNA sequence.  
 ACCESSION BES17162  
 VERSION BES17162.1 GI:9741192  
 KEYWORDS EST.  
 SOURCE bread wheat.  
 ORGANISM Triticum aestivum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae ; Triticeae; Triticum.  
 1 (bases 1 to 255)  
 Anderson,O.D., Chao,S., Han,P.S., Hsia,C.C., Johnson,R.R., Kang,Y., Lazo,G.R., Miller,R., Rausch,C.J., Seaton,C.L., Tong,J.C., Verhey ,S.D. and Walker-Simmons,M.K.  
 The structure and function of the expressed portion of the wheat genomes - ABA-treated embryo library  
 Unpublished (2000)  
 Contact: Olin Anderson  
 US Department of Agriculture, Agriculture Research Service, Pacific West Area, Western Regional Research Center

800 Buchanan Street, Albany, CA 94710, USA  
 Tel: 5105595773  
 Fax: 5105595818  
 Email: oanders@pw.usda.gov

Sequence have been trimmed to remove vector sequence and low quality sequence with phred score less than 20  
 Seq primer: Clontech Matchmaker 3' AD primer.

## FEATURES

source

1..265  
 /organism="Triticum aestivum"  
 /cultivar="Brevor (soft, white, winter, common wheat)"  
 /db\_xref="taxon:4565"  
 /clone="WHE0615\_H12\_P23"  
 /clone\_lib="Wheat ABA-treated embryo cDNA library"  
 /tissue\_type="Seed embryo"  
 /dev\_stage="Mature dormant seeds"  
 /lab\_host="E. coli DH12S"  
 /note="Vector: pCAD10; Site\_1: EcoRI; Site\_2: XhoI;  
 Embryos were cut from mature, dormant seeds and imbedded in 25 microm ABA (abscisic acid) in 5 mM Mes buffer, pH 5.7, for 12 hr at 22 C. The tissue, total RNA, and poly(A) RNA were prepared by Steven Verhey in M.K.  
 Walker-Simmons's lab (USDA-ARS, Washington State Univ., Pullman, Washington 99164-6420. A cDNA library was made by Clontech using a combination of random and oligo dt primers. Library was plated and archived by Russell Johnson (Colby College, ME/Walker-Simmons' lab). Plasmid DNA preparations and DNA sequencing were performed in the OD Anderson lab (all other authors)."  
 BASE COUNT 16 a 101 c 87 g 60 t 1 others  
 ORIGIN

Query Match 86.3%; Score 16.4; DB 10; Length 265;  
 Best Local Similarity 94.4%; Pred. No. 1e+03;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 2 CAAGGAGAGCGGTACGA 19  
 ||||| ||||| |||||  
 Db 169 CAAGGAGAGCGGTACGA 152

RESULT 11  
 AI948950/c  
 LOCUS  
 DEFINITION AI948950 290 bp mRNA linear EST 08-MAR-2000 wq18b12.xl NCI-CCAP\_Kid12 Homo sapiens cDNA clone IMAGE:2471615 3' similar to SW:GLI4\_HUMAN p10075 GLI4 PROTEIN ;, mRNA sequence.  
 ACCESSION AI948950  
 VERSION AI948950.1 GI:5741260  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 290)  
 NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index  
 Unpublished (1997)

REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 COMMENT  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-r@mail.nih.gov  
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.  
 cDNA Library Preparation: M. Bento Soares, Ph.D.  
 cDNA Library Arrayed by: Greg Lennon, Ph.D.  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CCAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html  
 Insert Length: 880 Std Error: 0.00  
 Seq primer: -40UP from Gibco  
 High quality sequence stop: 273.  
 FEATURES  
 source  
 1..290  
 Location/Qualifiers

/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2471615"  
/clone\_lib="NCI\_CGAP\_Kid12"  
/tissue\_type="2 pooled tumors (clear cell type)"  
/lab\_host="PH108"  
/note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site.1: Not I; Site.2: Eco RI; Plasmid DNA from the normalized library NCI\_CGAP\_Kid5 was prepared, and ss circles were made in vitro. Following HAP hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneIDs 1323912-1325831, 1471368-1472903 and 1492104-1493255). Subtraction by Bento Soares and M. Fatima Bonaldo." Subtraction by Bento Soares and M.  
Fatima Bonaldo

BASE COUNT 35 a 101 c 96 g 58 t  
ORIGIN

Query Match 86.3%; Score 16.4; DB 9; Length 290;  
Best Local Similarity 94.4%; Pred. No. 1e+03;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 CAAGGAGAGCGCTACGA 19  
|||||

Db 19 CAAGGAGAGCGCTACGA 2

RESULT 12  
BH648296  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

BH648296 375 bp DNA linear GSS 19-FEB-2002  
BOMFY48TF BO\_2\_3\_KB Brassica oleracea genomic clone BOMFY48, DNA sequence.  
BH648296  
BH648296.1 GI:18705920  
GSS.  
Brassica oleracea.  
Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.  
1 (bases 1 to 375)  
Town.C.D., Van Aken.S., Utterback.T. and Fraser.C.M.  
Whole genome shotgun sequencing of Brassica oleracea  
Unpublished (2001)  
Contact: Chris Town  
TIGR  
9712 Medical Center Drive, Rockville, MD 20850, USA.  
Tel: 301-838-3523  
Fax: 301-838-0208  
Email: cdtown@tigr.org  
DNA is from a doubled haploid provided by Tom Osborn.  
Seq primer: TF  
Class: sheared ends.

FEATURES  
source  
1..375  
Location/Qualifiers  
/organism="Brassica oleracea"  
/strain="T01000DH3"  
/db\_xref="taxon:3712"  
/clone="BOMFY48"  
/clone\_lib="BO\_2\_3\_KB"  
/note="Vector: pHOS1; Site.1: BstXI; 2-3 kb sheared genomic DNA inserted into pHOS1 using BstXI linkers"  
BASE COUNT 117 a 72 c 92 g 94 t  
ORIGIN

Query Match 86.3%; Score 16.4; DB 17; Length 375;  
Best Local Similarity 94.4%; Pred. No. 1.1e+03;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 CAAGGAGAGCGCTACGA 19  
|||||

Db 105 CAAGGAGAGCGCTACGA 122

RESULT 13  
H06751  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

H06751 432 bp mRNA linear EST 21-JUN-1995  
Y183a07.r1 Soares infant brain INIB Homo sapiens cDNA clone IMAGE:44512 5', mRNA sequence.  
H06751  
H06751.1 GI:870283  
EST.  
human.  
Homo sapiens  
Eukaryota: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevasakis,E., Waterston,R., Williamson,A., Woldmann,P. and Wilson,R.  
The WashU-Merck EST Project  
Unpublished (1995)  
Contact: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
Insert Size: 1388  
High quality sequence stops: 265  
Source: IMAGE Consortium, LLNL  
This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.  
Insert Length: 1388 Std Error: 0.00  
Seq primer: M13RP1  
High quality sequence stop: 265.  
Location/Qualifiers  
1..432  
/organism="Homo sapiens"  
/db\_xref="GDB:417053"  
/db\_xref="taxon:9606"  
/clone="IMAGE:44512"  
/clone\_lib="Soares infant brain INIB"  
/sex="female"  
/dev\_stage="73 days post natal"  
/lab\_host="DH10B (ampicillin resistant)"  
/note="Organ: whole brain; Vector: Lfamid BA; Site.1: Not I; Site.2: Hind III; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' AACTGGAAGAATTTCGCGCGCAGGAATTTTATTTT 3']; double-stranded cDNA was ligated to Hind III adaptors (Pharmacia), digested with Not I and directionally cloned into the Not I and Hind III sites of the Lfamid BA vector. Library went through one round of normalization. Library constructed by Bento Soares and M.Fatima Bonaldo."

BASE COUNT 98 a 101 c 102 g 125 t  
ORIGIN

Query Match 86.3%; Score 16.4; DB 14; Length 432;  
Best Local Similarity 94.4%; Pred. No. 1.2e+03;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 CAAGGAGAGCGCTACGA 19  
|||||

Db 335 CAAGGAGAGCGCTACAA 352

RESULT 14  
BM375873  
LOCUS  
DEFINITION  
ACCESSION  
VERSION

BM375873 437 bp mRNA linear EST 23-JUL-2002  
EBem06\_SQ004\_K13\_R embryo, 21 DPA, no treatment, cv Optic, EBem06 Hordeum vulgare cDNA clone EBem06\_SQ004\_K13 5', mRNA sequence.  
BM375873  
BM375873.2 GI:21935809

KEYWORDS EST.  
SOURCE Hordeum vulgare.  
ORGANISM Hordeum vulgare  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae  
; Triticeae; Hordeum.  
REFERENCE 1 (bases 1 to 437)  
AUTHORS Hedley, P., Liu, H., Caldwell, D., McCallum, N., Mudie, S., Cardle, L.,  
Ramsay, L., Machray, G., Marshall, D.F.M. and Waugh, R.  
TITLE Development of Barley Transcriptome Resources  
JOURNAL Unpublished (2001)  
COMMENT On Jan 10, 2002 this sequence version replaced gi:18119263.  
Contact: Waugh R, Marshall DF  
Genome Dynamics/Computational Biology  
Scottish Crop Research Institute  
Invergowrie, Dundee, DD2 5DA, Scotland, UK  
Tel: 00 44 1382 562731  
Fax: 00 44 1382 562426  
Email: est@scri.sari.ac.uk  
All sequence has a Phred quality score of 20 or over  
Seq primer: M13 reverse.

FEATURES Location/Qualifiers  
source  
1..437  
/organism="Hordeum vulgare"  
/cultivar="Optic"  
/db\_xref="taxon:4513"  
/clone="EBem06\_SQ004\_K13"  
/clone\_lib="embryo, 2l DPA, no treatment, cv Optic,  
EBem06"  
/tissue\_type="embryo"  
/dev\_stage="21 DPA"  
/lab\_host="DH108"  
/note="Vector: pSPORT1; Site\_1: Sal I; Site\_2: Not I;  
Non-normalised library, directionally cloned into pSPORT1.  
Derived from embryos dissected from developing grains (21  
days post anthesis) in glasshouse grown barley plants.  
Developed as part of the barley transcriptome resources of  
BBSRC/SEERAD funded cereal IGF (Investigating Gene  
Function) project."

BASE COUNT 97 a 138 c 154 g 48 t  
ORIGIN

Query Match 86.3%; Score 16.4; DB 13; Length 437;  
Best Local Similarity 94.4%; Pred. No. 1.2e+03;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 CAAGGAGAGCGCTACGA 19  
|||||||  
Db 345 CAAGGAGAGCGCTACGA 362

RESULT 15  
BH717250/c  
LOCUS BH717250 441 bp DNA linear GSS 20-FEB-2002  
DEFINITION BOMFM36TF BO\_2\_3\_KB Brassica oleracea genomic clone BOMFM36, DNA  
sequence.  
ACCESSION BH717250  
VERSION BH717250.1 GI:18814189  
KEYWORDS GSS.  
SOURCE Brassica oleracea.  
ORGANISM Brassica oleracea  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.  
REFERENCE 1 (bases 1 to 441)  
AUTHORS Town, C.D., Van Aken, S., Utterback, T. and Fraser, C.M.  
TITLE Whole genome shotgun sequencing of Brassica oleracea  
JOURNAL Unpublished (2001)  
COMMENT Other\_GSSs: BOMFM36TR  
Contact: Chris Town  
TIGR  
9712 Medical Center Drive, Rockville, MD 20850, USA.  
Tel: 301-838-3523

Fax: 301-838-0208  
Email: cdtown@tigr.org  
DNA is from a doubled haploid provided by Tom Osborn.  
Seq primer: TF  
Class: sheared ends.

FEATURES Location/Qualifiers  
source  
1..441  
/organism="Brassica oleracea"  
/strain="TO1000DH3"  
/db\_xref="taxon:3712"  
/clone="BOMFM36"  
/clone\_lib="BO\_2\_3\_KB"  
/note="Vector: pHOSt1; Site\_1: BstXI; 2-3 kb sheared  
genomic DNA inserted into pHOSt1 using BstXI linkers"

BASE COUNT 101 a 111 c 86 g 143 t  
ORIGIN

Query Match 86.3%; Score 16.4; DB 17; Length 441;  
Best Local Similarity 94.4%; Pred. No. 1.2e+03;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 CAAGGAGAGCGCTACGA 19  
|||||||  
Db 303 CAAGGAGAGCGCTACGA 286

Search completed: November 13, 2002, 04:01:02  
Job time : 114.815 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 12, 2002, 16:05:40 ; Search time 1.93295 Seconds  
(without alignments)  
3487.380 Million cell updates/sec

Title: US-09-697-123B-25  
Perfect score: 19  
Sequence: 1 tcaagagagcgtacga 19

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Capext 1.0

Searched: 320260 seqs, 177392727 residues

Total number of hits satisfying chosen parameters: 640520

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications, NA:\*

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*
- 12: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*
- 13: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*
- 14: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	16	84.2	212	10	US-09-833-381-29
2	15.8	83.2	265	10	US-09-294-093B-4362
3	15.8	83.2	281	10	US-09-294-093B-5081
4	15.8	83.2	1176	10	US-09-815-242-9077
5	15.8	83.2	6082	10	US-09-880-107-2432
6	15.4	81.1	331	10	US-09-860-352-9053
7	15.4	81.1	365	10	US-09-960-352-6002
8	15.4	81.1	370	10	US-09-960-352-9968
9	15.4	81.1	375	10	US-09-960-352-12378
10	15.4	81.1	384	10	US-09-960-352-8128
11	15.4	81.1	386	10	US-09-960-352-11656
12	15.4	81.1	390	10	US-09-960-352-5030
13	15.4	81.1	405	10	US-09-960-352-8824
14	15.4	81.1	414	10	US-09-960-352-14725
15	15.4	81.1	422	10	US-09-960-352-4589
16	15.4	81.1	1072	10	US-09-853-386-105
17	15.4	81.1	1230	10	US-09-960-288-1
18	15	78.9	1238	9	US-09-992-598-356
19	15	78.9	1238	10	US-09-989-722-356

20	15	78.9	1238	10	US-09-989-723-356	Sequence 356, App
21	15	78.9	1238	10	US-09-989-279-356	Sequence 356, App
22	15	78.9	1238	10	US-09-989-727-356	Sequence 356, App
23	15	78.9	1238	10	US-09-989-731-356	Sequence 356, App
24	15	78.9	1238	10	US-09-989-732-356	Sequence 356, App
25	15	78.9	1238	10	US-09-991-073-356	Sequence 356, App
26	15	78.9	1238	10	US-09-990-442-356	Sequence 356, App
27	15	78.9	1238	10	US-09-991-163-356	Sequence 356, App
28	15	78.9	1238	10	US-09-993-604-356	Sequence 356, App
29	15	78.9	1238	10	US-09-990-456-356	Sequence 356, App
30	15	78.9	1238	10	US-09-989-721-356	Sequence 356, App
31	14.8	77.9	143	10	US-09-983-965-2238	Sequence 2238, App
32	14.8	77.9	207	10	US-09-294-093B-3785	Sequence 3785, App
33	14.8	77.9	513	10	US-09-815-242-7608	Sequence 7608, App
34	14.8	77.9	520	10	US-09-867-701-6244	Sequence 6244, App
35	14.4	75.8	276	10	US-09-294-093B-4114	Sequence 4114, App
36	14.4	75.8	278	10	US-09-878-574-249	Sequence 249, App
37	14.4	75.8	289	10	US-09-960-352-3361	Sequence 3361, App
38	14.4	75.8	363	10	US-09-974-300-4785	Sequence 4785, App
39	14.2	74.7	278	10	US-09-923-876-4313	Sequence 4313, App
40	14.2	74.7	283	10	US-09-294-093B-6201	Sequence 6201, App
41	14.2	74.7	298	10	US-09-294-093B-4897	Sequence 4897, App
42	14.2	74.7	401	10	US-09-867-701-3765	Sequence 3765, App
43	14.2	74.7	415	10	US-09-960-352-1970	Sequence 1970, App
44	14.2	74.7	437	10	US-09-764-870-87	Sequence 87, Appl
45	14.2	74.7	588	10	US-09-867-550-1203	Sequence 1203, App

## ALIGNMENTS

RESULT 1  
US-09-833-381-29  
; Sequence 29, Application US/09833381  
; Patent No. US20020132090A1  
; GENERAL INFORMATION:  
; APPLICANT: Robison, Keith E.  
; TITLE OF INVENTION: No. US20020132090A1el Nucleic Acid and Protein Homologs  
; FILE REFERENCE: 5800-119  
; CURRENT APPLICATION NUMBER: US/09/833,381  
; CURRENT FILING DATE: 2001-04-11  
; PRIOR APPLICATION NUMBER: 09/516,448  
; PRIOR FILING DATE: 2000-02-29  
; NUMBER OF SEQ ID NOS: 2050  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 29  
; LENGTH: 212  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc.feature  
; LOCATION: (1)..(212)  
; OTHER INFORMATION: n - A,T,C or G  
US-09-833-381-29

Query Match 84.2%; Score 16; DB 10; Length 212;  
Best Local Similarity 84.2%; Pred. No. 20;  
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 1 TCAAGGAGAGCGCTACGA 19  
DB 75 TCANGAGAGCNCCTACGA 93

RESULT 2  
US-09-294-093B-4362  
; Sequence 4362, Application US/09294093B  
; Patent No. US20010051335A1  
; GENERAL INFORMATION:  
; APPLICANT: Lalgudi, Raghunath, V.  
; APPLICANT: Ito, Laura, Y.  
; APPLICANT: Sherman, Bradley, K.  
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN TASSEL

```
; FILE REFERENCE: PL-0009 US
; CURRENT APPLICATION NUMBER: US/09/294,093B
; CURRENT FILING DATE: 1999-04-16
; PRIOR APPLICATION NUMBER: 60/082,567
; PRIOR FILING DATE: April 21, 1998
; NUMBER OF SEQ ID NOS: 6207
; SOFTWARE: PERL Program
; SEQ ID NO 4362
; LENGTH: 265
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No. US20010051335A1 700353887H1
; NAME/KEY: unsure
; LOCATION: 69, 75, 80, 87, 206
; OTHER INFORMATION: a, t, c, g, or other
US-09-294-093B-4362

Query Match      83.2%; Score 15.8; DB 10; Length 265;
Best Local Similarity 89.5%; Pred. No. 25;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1 TCAAGGAGAGCGCTACGA 19
      |||||
Db      140 TCAAGGAGAGCGCTCCGA 158

RESULT 3
US-09-294-093B-5081
; Sequence 5081, Application US/09294093B
; Patent No. US20010051335A1
; GENERAL INFORMATION:
; APPLICANT: Ito, Laura, Y.
; APPLICANT: Sherman, Bradley, K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN TASSEL
; FILE REFERENCE: PL-0009 US
; CURRENT APPLICATION NUMBER: US/09/294,093B
; CURRENT FILING DATE: 1999-04-16
; PRIOR APPLICATION NUMBER: 60/082,567
; PRIOR FILING DATE: April 21, 1998
; NUMBER OF SEQ ID NOS: 6207
; SOFTWARE: PERL Program
; SEQ ID NO 5081
; LENGTH: 281
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No. US20010051335A1 700355754H1
; NAME/KEY: unsure
; LOCATION: 178, 205, 215
; OTHER INFORMATION: a, t, c, g, or other
US-09-294-093B-5081

Query Match      83.2%; Score 15.8; DB 10; Length 281;
Best Local Similarity 89.5%; Pred. No. 25;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1 TCAAGGAGAGCGCTACGA 19
      |||||
Db      243 TCAAGGAGAGCGCTCCGA 261

RESULT 4
US-09-815-242-9077
; Sequence 9077, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
```

```
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011a
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9077
; LENGTH: 1176
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1176)
US-09-815-242-9077

Query Match      83.2%; Score 15.8; DB 10; Length 1176;
Best Local Similarity 89.5%; Pred. No. 29;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1 TCAAGGAGAGCGCTACGA 19
      |||||
Db      131 TCAAGGAGAGCGCTACGA 149

RESULT 5
US-09-880-107-2432/c
; Sequence 2432, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2432
; LENGTH: 6082
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 M96233
US-09-880-107-2432

Query Match      83.2%; Score 15.8; DB 10; Length 6082;
Best Local Similarity 89.5%; Pred. No. 34;
```

```
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 TCAAGGAGAAGCGCTACGA 19
    ||||| ||||| ||| ||
Db 2645 TCAAGGAGAAGCGCTACGA 2627

RESULT 6
US-09-960-352-9053/c
; Sequence 9053, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 9053
; LENGTH: 331
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 39-LIB3057-003-Q1-K1-B4
US-09-960-352-9053

Query Match 81.1%; Score 15.4; DB 10; Length 331;
Best Local Similarity 94.1%; Pred. No. 41;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 CAAGGAGAAGCGCTACG 18
    ||||| ||||| |||||
Db 327 CAAGGAGAAGCGCTACG 311

RESULT 7
US-09-960-352-6002/c
; Sequence 6002, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 6002
; LENGTH: 365
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 26-LIB3057-001-Q1-K1-G9
US-09-960-352-6002

Query Match 81.1%; Score 15.4; DB 10; Length 365;
Best Local Similarity 94.1%; Pred. No. 42;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 CAAGGAGAAGCGCTACG 18
    ||||| ||||| |||||
Db 309 CAAGGAGAAGCGCTACG 293

RESULT 8
US-09-960-352-9968/c
; Sequence 9968, Application US/09960352
; Patent No. US20020137139A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 9968
; LENGTH: 370
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 43-LIB3057-006-Q1-K1-C12
US-09-960-352-9968

Query Match 81.1%; Score 15.4; DB 10; Length 370;
Best Local Similarity 94.1%; Pred. No. 42;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 CAAGGAGAAGCGCTACG 18
    ||||| ||||| |||||
Db 321 CAAGGAGAAGCGCTACG 305

RESULT 9
US-09-960-352-12378/c
; Sequence 12378, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 12378
; LENGTH: 375
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 53-LIB3057-014-Q1-K1-F10
US-09-960-352-12378

Query Match 81.1%; Score 15.4; DB 10; Length 375;
Best Local Similarity 94.1%; Pred. No. 42;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 CAAGGAGAAGCGCTACG 18
    ||||| ||||| |||||
Db 319 CAAGGAGAAGCGCTACG 303

RESULT 10
US-09-960-352-8128/c
; Sequence 8128, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
```

; SEQ ID NO 8128  
; LENGTH: 384  
; TYPE: DNA  
; ORGANISM: Bos taurus  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (9)  
; OTHER INFORMATION: unsure at all n locations  
; OTHER INFORMATION: Clone ID: 35-LIB3057-006-Q1-K1-A12  
US-09-960-352-8128

Query Match 81.1%; Score 15.4; DB 10; Length 384;  
Best Local Similarity 94.1%; Pred. No. 42;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CAAGGAGAGCGCTACG 18  
||||| |||||||  
Db 321 CAAGGACAAGCGCTACG 305

RESULT 11  
; Sequence 11656, Application US/09960352  
; Patent No. US20020137139A1  
; GENERAL INFORMATION:  
; APPLICANT: Warren, Wesley C.  
; APPLICANT: Tao, Nengbing  
; APPLICANT: Byatt, John C.  
; APPLICANT: Mathialagan, Nagappan  
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND  
; FILE REFERENCE: 16511.006/37-21(10298)C  
; CURRENT APPLICATION NUMBER: US/09/960,352  
; CURRENT FILING DATE: 2001-09-24  
; NUMBER OF SEQ ID NOS: 15112  
; SEQ ID NO 11656  
; LENGTH: 386  
; TYPE: DNA  
; ORGANISM: Bos taurus  
; OTHER INFORMATION: Clone ID: 50-LIB3057-003-Q1-K1-E10  
US-09-960-352-11656

Query Match 81.1%; Score 15.4; DB 10; Length 386;  
Best Local Similarity 94.1%; Pred. No. 42;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CAAGGAGAGCGCTACG 18  
||||| |||||||  
Db 201 CAAGGACAAGCGCTACG 217

RESULT 12  
; Sequence 5030, Application US/09960352  
; Patent No. US20020137139A1  
; GENERAL INFORMATION:  
; APPLICANT: Warren, Wesley C.  
; APPLICANT: Tao, Nengbing  
; APPLICANT: Byatt, John C.  
; APPLICANT: Mathialagan, Nagappan  
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND  
; FILE REFERENCE: 16511.006/37-21(10298)C  
; CURRENT APPLICATION NUMBER: US/09/960,352  
; CURRENT FILING DATE: 2001-09-24  
; NUMBER OF SEQ ID NOS: 15112  
; SEQ ID NO 5030  
; LENGTH: 390  
; TYPE: DNA  
; ORGANISM: Bos taurus  
; OTHER INFORMATION: Clone ID: 22-LIB3057-005-Q1-K1-F9  
US-09-960-352-5030

Query Match 81.1%; Score 15.4; DB 10; Length 390;  
Best Local Similarity 94.1%; Pred. No. 42;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CAAGGAGAGCGCTACG 18  
||||| |||||||  
Db 298 CAAGGACAAGCGCTACG 314

RESULT 13  
; Sequence 8824, Application US/09960352  
; Patent No. US20020137139A1  
; GENERAL INFORMATION:  
; APPLICANT: Warren, Wesley C.  
; APPLICANT: Tao, Nengbing  
; APPLICANT: Byatt, John C.  
; APPLICANT: Mathialagan, Nagappan  
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND  
; FILE REFERENCE: 16511.006/37-21(10298)C  
; CURRENT APPLICATION NUMBER: US/09/960,352  
; CURRENT FILING DATE: 2001-09-24  
; NUMBER OF SEQ ID NOS: 15112  
; SEQ ID NO 8824  
; LENGTH: 405  
; TYPE: DNA  
; ORGANISM: Bos taurus  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (383), (386), (392)  
; OTHER INFORMATION: unsure at all n locations  
; OTHER INFORMATION: Clone ID: 38-LIB3057-001-Q1-K1-B10  
US-09-960-352-8824

Query Match 81.1%; Score 15.4; DB 10; Length 405;  
Best Local Similarity 94.1%; Pred. No. 42;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CAAGGAGAGCGCTACG 18  
||||| |||||||  
Db 312 CAAGGACAAGCGCTACG 296

RESULT 14  
; Sequence 14725, Application US/09960352  
; Patent No. US20020137139A1  
; GENERAL INFORMATION:  
; APPLICANT: Warren, Wesley C.  
; APPLICANT: Tao, Nengbing  
; APPLICANT: Byatt, John C.  
; APPLICANT: Mathialagan, Nagappan  
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND  
; FILE REFERENCE: 16511.006/37-21(10298)C  
; CURRENT APPLICATION NUMBER: US/09/960,352  
; CURRENT FILING DATE: 2001-09-24  
; NUMBER OF SEQ ID NOS: 15112  
; SEQ ID NO 14725  
; LENGTH: 414  
; TYPE: DNA  
; ORGANISM: Bos taurus  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (346)  
; OTHER INFORMATION: unsure at all n locations  
; OTHER INFORMATION: Clone ID: 63-LIB3057-002-Q1-K1-H4  
US-09-960-352-14725

Query Match 81.1%; Score 15.4; DB 10; Length 414;  
Best Local Similarity 94.1%; Pred. No. 42;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CAAGGAGAGCGCTACG 18  
||||| |||||||||  
Db 167 CAAGGACAAGCGCTACG 183

## RESULT 15

US-09-960-352-4589/c  
; Sequence 4589, Application US/09960352  
; Patent No. US20020137139A1  
; GENERAL INFORMATION:  
; APPLICANT: Warren, Wesley C.  
; APPLICANT: Tao, Nengbing  
; APPLICANT: Byatt, John C.  
; APPLICANT: Mathialagan, Nagappan  
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND  
; FILE REFERENCE: 16511.006/37-21(10298)C  
; CURRENT APPLICATION NUMBER: US/09/960,352  
; CURRENT FILING DATE: 2001-09-24  
; NUMBER OF SEQ ID NOS: 15112  
; SEQ ID NO 4589  
; LENGTH: 422  
; TYPE: DNA  
; ORGANISM: Bos taurus  
; OTHER INFORMATION: Clone ID: 20-LIB3057-020-Q1-K1-E11  
US-09-960-352-4589

Query Match 81.1%; Score 15.4; DB 10; Length 422;  
Best Local Similarity 94.1%; Pred. No. 42;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CAAGGAGAGCGCTACG 18  
||||| |||||||||  
Db 312 CAAGGACAAGCGCTACG 296

Search completed: November 12, 2002, 16:59:30  
Job time : 4.93295 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 12, 2002, 15:56:36 : Search time 2.54798 Seconds  
(without alignments)  
286.857 Million cell updates/sec

Title: US-09-697-123B-25

Perfect score: 19

Sequence: 1 tcaaggagaagcgctacga 19

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued\_Patents\_NA.\*  
1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PTUS\_COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	19	100.0	970	1	US-08-250-030-1
2	19	100.0	970	5	PCT-US95-06790-1
3	17.4	91.6	3447	2	US-08-313-185-57
4	17.4	91.6	3447	3	US-08-082-614A-57
5	15.8	83.2	1958	4	US-08-961-527-360
6	15.8	83.2	3752	4	US-08-961-527-208
7	15.8	83.2	4854	4	US-08-961-527-178
8	15.8	83.2	6516	4	US-08-961-527-105
9	15.8	83.2	11309	4	US-08-961-527-108
10	14.8	77.9	2395	4	US-09-724-864-24
11	14.8	77.9	6559	4	US-09-234-186-1
12	14.8	77.9	6559	4	US-09-233-527-1
13	14.8	77.9	6560	5	PCT-US93-05651-1
14	14.4	75.8	515	4	US-08-998-416-125
15	14.4	75.8	711	4	US-09-221-017B-476
16	14.2	74.7	22	2	US-08-494-151-6
17	14.2	74.7	756	4	US-08-413-974-3
18	14.2	74.7	756	4	US-08-434-418-3
19	14.2	74.7	756	4	US-08-433-288-3
20	14.2	74.7	756	4	US-08-174-739A-3
21	14.2	74.7	756	4	US-08-434-256-3
22	14.2	74.7	805	4	US-08-858-207A-1
23	14.2	74.7	810	4	US-08-413-974-5
24	14.2	74.7	810	4	US-08-434-418-5
25	14.2	74.7	810	4	US-08-433-288-5
26	14.2	74.7	810	4	US-08-174-739A-5
27	14.2	74.7	810	4	US-08-434-256-5

28	14.2	74.7	812	4	US-09-091-097-7	Sequence 7, Appli
29	14.2	74.7	1049	4	US-08-961-527-317	Sequence 317, App
30	14.2	74.7	1123	1	US-07-971-096-3	Sequence 3, Appli
31	14.2	74.7	1123	1	US-08-175-096-3	Sequence 3, Appli
32	14.2	74.7	2091	4	US-09-134-001C-1459	Sequence 1459, App
33	14.2	74.7	2495	4	US-08-961-527-255	Sequence 255, App
34	14.2	74.7	3240	4	US-09-171-337A-4	Sequence 4, Appli
35	14.2	74.7	4320	2	US-08-494-151-13	Sequence 13, Appli
36	14.2	74.7	4695	2	US-08-231-193A-57	Sequence 57, Appli
37	14.2	74.7	4695	2	US-08-486-273A-57	Sequence 57, Appli
38	14.2	74.7	4695	3	US-08-940-086A-57	Sequence 57, Appli
39	14.2	74.7	4695	4	US-08-940-035A-57	Sequence 57, Appli
40	14.2	74.7	4695	4	US-08-935-105A-57	Sequence 57, Appli
41	14.2	74.7	4695	4	US-09-648-797-57	Sequence 57, Appli
42	14.2	74.7	6171	4	US-08-961-527-37	Sequence 37, Appli
43	14.2	74.7	7650	4	US-09-221-017B-911	Sequence 911, App
44	14.2	74.7	8948	4	US-09-643-597-119	Sequence 119, App
45	14	73.7	7301	4	US-09-816-088-3	Sequence 3, Appli

#### ALIGNMENTS

RESULT 1  
US-08-250-030-1  
: Sequence 1, Application US/08250030  
: Patent No. 5643723  
: GENERAL INFORMATION:  
: APPLICANT: Persing, David H.  
: TITLE OF INVENTION: Detection of a Genetic Locus Encoding  
: TITLE OF INVENTION: Resistance to Rifampin in Mycobacterial Cultures and in  
: NUMBER OF SEQUENCES: 15  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: Schwegman, Lundberg & Woessner  
: STREET: 3500 IDS Center  
: CITY: Minneapolis  
: STATE: MN  
: COUNTRY: USA  
: ZIP: 55402  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: COMPUTER: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: PatentIn Release #1.0, Version #1.25  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/08/250,030  
: FILING DATE: 26-MAY-1994  
: CLASSIFICATION: 435  
: ATTORNEY/AGENT INFORMATION:  
: NAME: Muetting, Ann M.  
: REGISTRATION NUMBER: 33,977  
: REFERENCE/DOCKET NUMBER: 150.105U1  
: TELEPHONE: 612-339-0331  
: TELEFAX: 612-339-3061  
: INFORMATION FOR SEQ ID NO: 1:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 970 base pairs  
: TYPE: nucleic acid  
: STRANDEDNESS: single  
: TOPOLOGY: linear  
: MOLECULE TYPE: DNA  
US-08-250-030-1

Query Match 100.0%; Score 19; DB 1; Length 970;  
Best Local Similarity 100.0%; Pred. No. 1.2;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCAAGGAGAAGCGCTACGA 19  
Db 26 TCAAGGAGAAGCGCTACGA 44

RESULT 2  
PCT-US95-06790-1  
; Sequence 1, Application PC/TUS9506790  
; GENERAL INFORMATION:  
; APPLICANT: Mayo Foundation for Medical Education and Research  
; APPLICANT: and Hoffmann-La Roche Inc.  
; TITLE OF INVENTION: Detection of a Genetic Locus Encoding  
; TITLE OF INVENTION: Resistance to Rifampin  
; NUMBER OF SEQUENCES: 15  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Schwegman, Lundberg & Woessner  
; STREET: 3500 IDS Center  
; CITY: Minneapolis  
; STATE: MN  
; COUNTRY: USA  
; ZIP: 55402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/06790  
; FILING DATE: 26-MAY-1995  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Raasch, Kevin W.  
; REGISTRATION NUMBER: 35,651  
; REFERENCE/DOCKET NUMBER: 150.105W01  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 612-339-0331  
; TELEFAX: 612-339-3061  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 970 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA  
PCT-US95-06790-1

Query Match 100.0%; Score 19; DB 5; Length 970;  
Best Local Similarity 100.0%; Pred. No. 1.2;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCAAGGAGAGCGCTACGA 19  
Db 26 TCAAGGAGAGCGCTACGA 44

RESULT 3  
US-08-313-185-57  
; Sequence 57, Application US/08313185  
; Patent No. 5851763  
; GENERAL INFORMATION:  
; APPLICANT: Heym, Beate  
; APPLICANT: Cole, Stewart  
; APPLICANT: Young, Douglas  
; APPLICANT: Zhang, Ying  
; APPLICANT: Honore, Nadine  
; APPLICANT: Telenti, Amalio  
; APPLICANT: Bodmer, Thomas  
; TITLE OF INVENTION: Rapid Detection of Antibiotic Resistance  
; TITLE OF INVENTION: In Mycobacterium Tuberculosis  
; NUMBER OF SEQUENCES: 66  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &  
; STREET: 1300 I Street, N.W.  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA

ZIP: 20005-3315  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/313,185  
; FILING DATE: 12-OCT-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Meyers, Kenneth J.  
; REGISTRATION NUMBER: 25,146  
; REFERENCE/DOCKET NUMBER: 02356.0068-00000  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 408-4000  
; TELEFAX: (202) 408-4400  
; INFORMATION FOR SEQ ID NO: 57:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3447 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
US-08-313-185-57

Query Match 91.6%; Score 17.4; DB 2; Length 3447;  
Best Local Similarity 94.7%; Pred. No. 8.2;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TCAAGGAGAGCGCTACGA 19  
Db 809 TCAAGGAGAGCGCTACGA 827

RESULT 4  
US-09-082-614A-57  
; Sequence 57, Application US/09082614A  
; Patent No. 6124098  
; GENERAL INFORMATION:  
; APPLICANT: Heym, Beate  
; APPLICANT: Cole, Stewart  
; APPLICANT: Young, Douglas  
; APPLICANT: Zhang, Ying  
; APPLICANT: Honore, Nadine  
; APPLICANT: Telenti, Amalio  
; APPLICANT: Bodmer, Thomas  
; TITLE OF INVENTION: Rapid Detection of Antibiotic Resistance  
; TITLE OF INVENTION: In Mycobacterium Tuberculosis  
; NUMBER OF SEQUENCES: 66  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &  
; STREET: 1300 I Street, N.W.  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20005-3315  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/082,614A  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/313,185  
; FILING DATE: 12-OCT-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Meyers, Kenneth J.  
; REGISTRATION NUMBER: 25,146

REFERENCE/DOCKET NUMBER: 02356.0068-00000  
TELEPHONE: (202) 408-4000  
TELEFAX: (202) 408-4400  
INFORMATION FOR SEQ ID NO: 57:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3447 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-09-082-614A-57

Query Match 91.6%; Score 17.4; DB 3; Length 3447;  
Best Local Similarity 94.7%; Pred. No. 8.2;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TCAAGGAGAGCGCTACGA 19  
||||| |||||

Db 809 TCAAGGAGAGCGCTACGA 827

## RESULT 5

US-08-961-527-360  
Sequence 360, Application US/08961527  
Patent No. 6420135  
GENERAL INFORMATION:

APPLICANT: Charles Kunsch  
TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences  
NUMBER OF SEQUENCES: 391  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: Maryland  
COUNTRY: USA  
ZIP: 20850

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
COMPUTER: HP Vectra 486/33  
OPERATING SYSTEM: MSDOS version 6.2  
SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/961,527  
FILING DATE:

PRIOR APPLICATION DATA:  
CLASSIFICATION: 424  
APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Brookes, A. Anders  
REGISTRATION NUMBER: 36,373

REFERENCE/DOCKET NUMBER: PB340P1

TELEPHONE: (301) 309-8504

TELEFAX: (301) 309-8512

INFORMATION FOR SEQ ID NO: 360:

SEQUENCE CHARACTERISTICS:

LENGTH: 1958 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

US-08-961-527-360

Query Match 83.2%; Score 15.8; DB 4; Length 1958;  
Best Local Similarity 89.5%; Pred. No. 46;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 TCAAGGAGAGCGCTACGA 19

||||| |||||

Db 194 TCAAGGAGAGCGCTACGA 212

## RESULT 6

US-08-961-527-208  
Sequence 208, Application US/08961527  
Patent No. 6420135  
GENERAL INFORMATION:

APPLICANT: Charles Kunsch

TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences

NUMBER OF SEQUENCES: 391

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville

STATE: Maryland

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage

COMPUTER: HP Vectra 486/33

OPERATING SYSTEM: MSDOS version 6.2

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/961,527

FILING DATE:

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Brookes, A. Anders

REGISTRATION NUMBER: 36,373

REFERENCE/DOCKET NUMBER: PB340P1

TELEPHONE: (301) 309-8504

TELEFAX: (301) 309-8512

INFORMATION FOR SEQ ID NO: 208:

SEQUENCE CHARACTERISTICS:

LENGTH: 3752 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

US-08-961-527-208

Query Match 83.2%; Score 15.8; DB 4; Length 3752;  
Best Local Similarity 89.5%; Pred. No. 49;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 TCAAGGAGAGCGCTACGA 19

||||| |||||

Db 2773 TCAAGGAGAGCGCTACGA 2791

## RESULT 7

US-08-961-527-178  
Sequence 178, Application US/08961527  
Patent No. 6420135  
GENERAL INFORMATION:

APPLICANT: Charles Kunsch

TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences

NUMBER OF SEQUENCES: 391

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville

STATE: Maryland

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage

COMPUTER: HP Vectra 486/33

OPERATING SYSTEM: MSDOS version 6.2

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/961,527



; FILING DATE:  
 ; CLASSIFICATION: 424  
 ; PRIOR APPLICATION DATA:  
 ; FILING DATE:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Brookes, A. Anders  
 ; REGISTRATION NUMBER: 36,373  
 ; REFERENCE/DOCKET NUMBER: PB340P1  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (301) 309-8504  
 ; INFORMATION FOR SEQ ID NO: 178:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 4854 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: double  
 ; TOPOLOGY: linear  
 ; US-08-961-527-178

Query Match 83.2%; Score 15.8; DB 4; Length 4854;  
 Best Local Similarity 89.5%; Pred. No. 50;  
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TCAAGGAGGAGCGCTACGA 19  
 ||||| ||||| |||||  
 DB 100 TCAAAGAGAGCACTACGA 118

RESULT 8  
 US-08-961-527-105/c  
 ; Sequence 105, Application US/08961527  
 ; Patent No. 6420135  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Charles Kunsch  
 ; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences  
 ; NUMBER OF SEQUENCES: 391  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Human Genome Sciences, Inc.  
 ; STREET: 9410 Key West Avenue  
 ; CITY: Rockville  
 ; STATE: Maryland  
 ; COUNTRY: USA  
 ; ZIP: 20850  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
 ; COMPUTER: HP Vectra 486/33  
 ; OPERATING SYSTEM: MSDOS version 6.2  
 ; SOFTWARE: ASCII Text  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/961,527  
 ; FILING DATE:  
 ; CLASSIFICATION: 424  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER:  
 ; FILING DATE:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Brookes, A. Anders  
 ; REGISTRATION NUMBER: 36,373  
 ; REFERENCE/DOCKET NUMBER: PB340P1  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (301) 309-8504  
 ; TELEFAX: (301) 309-8512  
 ; INFORMATION FOR SEQ ID NO: 105:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 6516 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: double  
 ; TOPOLOGY: linear  
 ; US-08-961-527-105

Query Match 83.2%; Score 15.8; DB 4; Length 6516;  
 Best Local Similarity 89.5%; Pred. No. 52;

Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 TCAAGGAGGAGCGCTACGA 19  
 ||||| ||||| |||||  
 DB 5353 TCAAAGAGAGCACTACGA 5335

RESULT 9  
 US-08-961-527-108/c  
 ; Sequence 108, Application US/08961527  
 ; Patent No. 6420135  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Charles Kunsch  
 ; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences  
 ; NUMBER OF SEQUENCES: 391  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Human Genome Sciences, Inc.  
 ; STREET: 9410 Key West Avenue  
 ; CITY: Rockville  
 ; STATE: Maryland  
 ; COUNTRY: USA  
 ; ZIP: 20850  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
 ; COMPUTER: HP Vectra 486/33  
 ; OPERATING SYSTEM: MSDOS version 6.2  
 ; SOFTWARE: ASCII Text  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/961,527  
 ; FILING DATE:  
 ; CLASSIFICATION: 424  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER:  
 ; FILING DATE:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Brookes, A. Anders  
 ; REGISTRATION NUMBER: 36,373  
 ; REFERENCE/DOCKET NUMBER: PB340P1  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (301) 309-8504  
 ; TELEFAX: (301) 309-8512  
 ; INFORMATION FOR SEQ ID NO: 108:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 11309 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: double  
 ; TOPOLOGY: linear  
 ; US-08-961-527-108

Query Match 83.2%; Score 15.8; DB 4; Length 11309;  
 Best Local Similarity 89.5%; Pred. No. 55;  
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TCAAGGAGGAGCGCTACGA 19  
 ||||| ||||| |||||  
 DB 9043 TCAAAGAGAGCACTACGA 9025

RESULT 10  
 US-09-724-864-24/c  
 ; Sequence 24, Application US/09724864  
 ; Patent No. 6380362  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Watson, James D  
 ; APPLICANT: Murison, James G.  
 ; TITLE OF INVENTION: Polynucleotides, polypeptides expressed  
 ; FILE REFERENCE: 11000.1050u1  
 ; CURRENT APPLICATION NUMBER: US/09/724,864  
 ; CURRENT FILING DATE: 2000-11-28  
 ; PRIOR APPLICATION NUMBER: U.S. No. 6380362 60/171,678  
 ; PRIOR FILING DATE: 1999-12-23  
 ; NUMBER OF SEQ ID NOS: 72

```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24
; LENGTH: 2395
; TYPE: DNA
; ORGANISM: Mouse
US-09-724-864-24

Query Match          77.9%; Score 14.8; DB 4; Length 2395;
Best Local Similarity 88.9%; Pred. No. 1.4e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy  2 CAAGGAGAGCGCTACGA 19
    |||||  |||||
Db  1796 CAAGGAGAACTCTACGA 1779

RESULT 11
US-09-234-186-1
; Sequence 1, Application US/09234186
; Patent No. 6312947
; GENERAL INFORMATION:
; APPLICANT: Horvitz, H. Robert
; TITLE OF INVENTION: IDENTIFICATION AND CHARACTERIZATION OF A
; TITLE OF INVENTION: GENE WHICH PROTECTS CELLS FROM PROGRAMMED CELL DEATH AND
; FILE REFERENCE: 01997/201005
; CURRENT APPLICATION NUMBER: US/09/234,186
; CURRENT FILING DATE: 1999-01-20
; EARLIER APPLICATION NUMBER: 07/898,933
; EARLIER FILING DATE: 1992-06-12
; EARLIER APPLICATION NUMBER: 07/927,681
; EARLIER FILING DATE: 1992-08-10
; EARLIER APPLICATION NUMBER: 08/288,295
; EARLIER FILING DATE: 1994-08-10
; EARLIER APPLICATION NUMBER: 08/801,248
; EARLIER FILING DATE: 1997-02-19
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 6559
; TYPE: DNA
; ORGANISM: Caenorhabditis elegans
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(6559)
; OTHER INFORMATION: n = A,T,C or G
US-09-234-186-1

Query Match          77.9%; Score 14.8; DB 4; Length 6559;
Best Local Similarity 88.9%; Pred. No. 1.6e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy  2 CAAGGAGAGCGCTACGA 19
    |||||  |||||
Db  5725 CAAGGAGAGTCTAAGA 5742

RESULT 12
US-09-233-527-1
; Sequence 1, Application US/09233527
; Patent No. 6465617
; GENERAL INFORMATION:
; APPLICANT: Horvitz, H. Robert
; TITLE OF INVENTION: IDENTIFICATION AND CHARACTERIZATION OF A
; TITLE OF INVENTION: GENE WHICH PROTECTS CELLS FROM PROGRAMMED CELL DEATH AND
; FILE REFERENCE: 01997/201004
; CURRENT APPLICATION NUMBER: US/09/233,527
; CURRENT FILING DATE: 1999-01-20
; PRIOR APPLICATION NUMBER: 07/898,933
; PRIOR FILING DATE: 1992-06-12
```

```
; PRIOR APPLICATION NUMBER: 07/927,681
; PRIOR FILING DATE: 1992-08-10
; PRIOR APPLICATION NUMBER: 08/288,295
; PRIOR FILING DATE: 1994-08-10
; PRIOR APPLICATION NUMBER: 08/801,248
; PRIOR FILING DATE: 1997-02-19
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 6559
; TYPE: DNA
; ORGANISM: Caenorhabditis elegans
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(6559)
; OTHER INFORMATION: n = A,T,C or G
US-09-233-527-1

Query Match          77.9%; Score 14.8; DB 4; Length 6559;
Best Local Similarity 88.9%; Pred. No. 1.6e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy  2 CAAGGAGAGCGCTACGA 19
    |||||  |||||
Db  5725 CAAGGAGAGTCTAAGA 5742

RESULT 13
PCT-US93-05651-1
; Sequence 1, Application PC/TUS9305651
; GENERAL INFORMATION:
; TITLE OF INVENTION: A Gene Which Prevents Programmed Cell Death
; NUMBER OF SEQUENCES: 5
; COMPUTER READABLE FORM:
; MEDIUM TYPE: diskette
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/05651
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6560 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
PCT-US93-05651-1

Query Match          77.9%; Score 14.8; DB 5; Length 6560;
Best Local Similarity 88.9%; Pred. No. 1.6e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy  2 CAAGGAGAGCGCTACGA 19
    |||||  |||||
Db  5726 CAAGGAGAGTCTAAGJ 5743

RESULT 14
US-08-998-416-125/c
; Sequence 125, Application US/08998416
; Patent No. 6239264
; GENERAL INFORMATION:
; APPLICANT: Philippsen, Peter
; APPLICANT: Pohlmann, Rainer
; APPLICANT: Steiner, Sabine
; APPLICANT: Mohr, Christine
; APPLICANT: Wendland, Jurgin
; APPLICANT: Knechtle, Philipp
; APPLICANT: Rebschung, Corinne
; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPPII
; TITLE OF INVENTION: AND USES THEREOF
; NUMBER OF SEQUENCES: 1152
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6239264artis Corporation
; STREET: 3054 Cornwallis Road
```

```
; CITY: Research Triangle Park
; STATE: NO 6239264th Carolina
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/998,416
; FILING DATE: 24-DEC-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: CH 0016/97
; FILING DATE: 31-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Meligs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: Pf/5-30306/A/GGC1976
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8587
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 125:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 515 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORGANISM: PAG1046RP
; US-08-998-416-125

Query Match 75.8%; Score 14.4; DB 4; Length 515;
Best Local Similarity 93.8%; Pred. No. 1.9e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 CAAGGAGAGCGCTAC 17
Db 415 CAAGGAGAGCGCTAC 400
||||| |||||

RESULT 15
US-09-221-017B-476/c
; Sequence 476, Application US/09221017B
; Patent No. 6444799
; GENERAL INFORMATION:
; APPLICANT: Ross, Bruce C.
; TITLE OF INVENTION: P. GINGIVALIS NUCLEOTIDES AND USES THEREOF
; NUMBER OF SEQUENCES: 1120
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 PAGE MILL ROAD
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/221,017B
; FILING DATE: 23-DEC-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PP1182
; FILING DATE: 31-DEC-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PP1546
; FILING DATE: 30-JAN-1998
```

```
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PP2911
; FILING DATE: 09-APR-1998
; PRIOR APPLICATION DATA: PCT/AU98/01023
; APPLICATION NUMBER:
; FILING DATE: 10-DEC-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Monroy, Gladys H
; REGISTRATION NUMBER: 32,430
; REFERENCE/DOCKET NUMBER: 27340-20021.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-813-5600
; TELEFAX: 650-494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 476:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 711 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: UNKNOWN
; ORIGINAL SOURCE:
; ORGANISM: PORYPHYROMONAS GINGIVALIS
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1...711
; US-09-221-017B-476

Query Match 75.8%; Score 14.4; DB 4; Length 711;
Best Local Similarity 93.8%; Pred. No. 2e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 CAAGGAGAGCGCTAC 17
Db 217 CAAGGAGAGCGCTAC 202
||||| ||||| |||||

Search completed: November 12, 2002, 21:03:46
Job time : 8.54798 secs
```

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 12, 2002, 15:44:35 ; Search time 12.4434 Seconds  
(without alignments)  
3438.621 Million cell updates/sec

Title: US-09-697-123B-25

Perfect score: 19

Sequence: 1 tcaaggagaagcgctacga 19

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N\_Geneseq\_101002:\*

1: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:\*

2: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:\*

3: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:\*

4: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:\*

5: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:\*

6: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:\*

7: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:\*

8: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:\*

9: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:\*

10: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:\*

11: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:\*

12: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:\*

13: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:\*

14: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:\*

15: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:\*

16: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:\*

17: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:\*

18: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:\*

19: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:\*

20: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:\*

21: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:\*

22: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:\*

23: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:\*

24: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	19	100.0	19	22	AAS05225 Mycobacterium spec
2	19	100.0	207	22	AAS05204 Mycobacterium gord
3	19	100.0	207	22	AAS05212 Mycobacterium szul
4	19	100.0	208	22	AAS05201 Mycobacterium gord
5	19	100.0	208	22	AAS05202 Mycobacterium gord
6	19	100.0	208	22	AAS05203 Mycobacterium gord
7	19	100.0	208	22	AAS05205 Mycobacterium tube
8	19	100.0	208	22	AAS05206 Mycobacterium terr
9	19	100.0	208	22	AAS05208 Mycobacterium kans

10	19	100.0	208	22	AAS05210 Mycobacterium ulce
11	19	100.0	208	22	AAS05211 Mycobacterium mari
12	19	100.0	208	22	AAS05215 Mycobacterium aviu
13	19	100.0	208	22	AAS05216 Mycobacterium bowi
14	19	100.0	208	22	AAS05217 Mycobacterium cela
15	19	100.0	208	22	AAS05219 Mycobacterium fort
16	19	100.0	208	22	AAS05222 Mycobacterium afri
17	19	100.0	208	22	AAS05223 Mycobacterium haem
18	19	100.0	208	22	AAS05224 Mycobacterium xeno
19	19	100.0	211	22	AAS05218 Mycobacterium flav
20	19	100.0	214	22	AAS05207 Mycobacterium chel
21	19	100.0	214	22	AAS05214 Mycobacterium mal
22	19	100.0	214	22	AAS05221 Mycobacterium abs
23	19	100.0	223	22	AAS05209 Mycobacterium scro
24	19	100.0	223	22	AAS05213 Mycobacterium gast
25	19	100.0	223	22	AAS05213 Mycobacterium tube
26	19	100.0	3519	22	AAT09676 Mycobacterium tube
27	19	100.0	3534	22	AAT09676 Mycobacterium tube
28	19	100.0	3853	21	AAH02079 Mycobacterium tube
29	19	100.0	3853	21	AAA74651 Mycobacterium tube
30	17.4	91.6	3239	23	ABL08943 M. tuberculosis rp
31	17.4	91.6	3239	23	ABL08943 Drosophila melanog
32	17.4	91.6	3447	14	AAQ51532 M.leprae rpoB gene
33	17.4	91.6	4525	23	ABL08942 Drosophila melanog
34	17.4	91.6	6144	23	ABL03424 Drosophila melanog
35	16.4	86.3	407	22	ABA20983 Human nervous syst
36	16.4	86.3	407	22	AAK89467 Human digestive sy
37	16.4	86.3	13644	23	ABL04469 Drosophila melanog
38	16.4	86.3	17549	23	ABL04468 Drosophila melanog
39	15.8	83.2	265	24	ABL74988 Corn tassal-derive
40	15.8	83.2	261	24	ABL75707 Corn tassal-derive
41	15.8	83.2	1040	19	AAV65213 DNA encoding a S.
42	15.8	83.2	1176	23	AAS55440 Streptococcus pneu
43	15.8	83.2	1958	19	AAV52493 Streptococcus pneu
44	15.8	83.2	2102	23	ABL10513 Drosophila melanog
45	15.8	83.2	3752	19	AAV52341 Streptococcus pneu

ALIGNMENTS

RESULT 1

AAS05225

ID AAS05225 standard; DNA; 19 BP.

XX AAS05225;

AC AAS05225;

XX 07-SEP-2001 (first entry)

DT 07-SEP-2001 (first entry)

XX Mycobacterium species rpoB gene fragment PCR primer RPO5'.

DE Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;

XX PCR-restriction fragment length polymorphism analysis; PCR primer; 58.

KW Mycobacterium sp.

XX WO200131061-A1.

PN 03-MAY-2001.

XX 27-OCT-2000; 2000WO-KR01223.

PF 27-OCT-1999; 99KR-0046795.

XX (ERUM-) ERUME BIOTECH CO LTD.

PA Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;

XX WPI; 2001-300520/31.

XX New DNA fragments from the rpoB gene of mycobacteria, useful for

PT diagnosis and identification of many mycobacterial species by

PT restriction fragment length polymorphism -

XX Disclosure; Page 13; 50pp; English.

XX The present sequence for Mycobacterium species rpoB gene fragment PCR  
XX primer RPO5' is used with PCR primer RPO3' (AAS05226) to amplify the  
CC rpoB gene region from 24 rpoB gene fragments (AAS05201-AAS05224) from  
CC various Mycobacterial species. These rpoB gene fragments can be used  
CC in the diagnosis and identification of Mycobacterium species using a  
CC novel PCR-restriction fragment length polymorphism analysis (PRA)  
CC method. The method comprises obtaining a restriction fragment length  
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,  
CC amplifying and digesting the DNA fragment from the microorganism to  
CC be identified and comparing the RFLP patterns from the known rpoB gene  
CC fragments with the unidentified fragment. The rpoB gene fragments  
CC are useful to identify a wide range of Mycobacterium species, e.g. for  
CC diagnosis or to obtain epidemiological and pathogenesis information for  
CC selection of appropriate therapies, including M. tuberculosis, M. leprae  
CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected  
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene  
CC fragments is rapid, precise, simple and cost effective (only 1 PCR  
CC required), and can differentiate between many species in a single  
CC experiment, including those difficult to distinguish by usual biochemical  
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for  
CC detecting specific Mycobacterial species.

SO Sequence 19 BP; 7 A; 4 C; 6 G; 2 T; 0 other;

Query Match 100.0%; Score 19; DB 22; Length 19;

Best Local Similarity 100.0%; Pred. No. 6.5; Mismatches 0; Indels 0; Gaps 0;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TCAGAGGAGAGCGCTACGA 19

DB 1 TCAGAGGAGAGCGCTACGA 19

RESULT 2  
AAS05204

ID AAS05204 standard; DNA; 207 BP.

AC AAS05204;

DT 07-SEP-2001 (first entry)

DE Mycobacterium gordonae type IV rpoB gene fragment.

KW Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;

KW PCR-restriction fragment length polymorphism analysis; ds.

OS Mycobacterium gordonae type IV.

PN WO200131061-A1.

PD 03-MAY-2001.

PF 27-OCT-2000; 2000WO-KR01223.

PR 27-OCT-1999; 99KR-0046795.

PA (ERUM-) ERUME BIOTECH CO LTD.

PI Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;

DR WPI; 2001-300520/31.

PT New DNA fragments from the rpoB gene of mycobacteria, useful for  
PT diagnosis and identification of many mycobacterial species by  
PT restriction fragment length polymorphism

PS Claim 1; Page 41; 50pp; English.

CC The present sequence for Mycobacterium gordonae type IV rpoB gene  
CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from

CC various Mycobacterial species. These rpoB gene fragments can be used  
CC in the diagnosis and identification of Mycobacterium species using a  
CC novel PCR-restriction fragment length polymorphism analysis (PRA)  
CC method. The method comprises obtaining a restriction fragment length  
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,  
CC amplifying and digesting the DNA fragment from the microorganism to  
CC be identified and comparing the RFLP patterns from the known rpoB gene  
CC fragments with the unidentified fragment. The rpoB gene fragments  
CC are useful to identify a wide range of Mycobacterium species, e.g. for  
CC diagnosis or to obtain epidemiological and pathogenesis information for  
CC selection of appropriate therapies, including M. tuberculosis, M. leprae  
CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected  
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene  
CC fragments is rapid, precise, simple and cost effective (only 1 PCR  
CC required), and can differentiate between many species in a single  
CC experiment, including those difficult to distinguish by usual biochemical  
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for  
CC detecting specific Mycobacterial species.

SO Sequence 207 BP; 45 A; 66 C; 66 G; 30 T; 0 other;

Query Match 100.0%; Score 19; DB 22; Length 207;

Best Local Similarity 100.0%; Pred. No. 7.5; Mismatches 0; Indels 0; Gaps 0;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TCAGAGGAGAGCGCTACGA 19

DB 1 TCAGAGGAGAGCGCTACGA 19

RESULT 3  
AAS05212

ID AAS05212 standard; DNA; 207 BP.

AC AAS05212;

DT 07-SEP-2001 (first entry)

DE Mycobacterium szulgai rpoB gene fragment.

KW Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;

KW PCR-restriction fragment length polymorphism analysis; ds.

OS Mycobacterium szulgai.

PN WO200131061-A1.

PD 03-MAY-2001.

PF 27-OCT-2000; 2000WO-KR01223.

PR 27-OCT-1999; 99KR-0046795.

PA (ERUM-) ERUME BIOTECH CO LTD.

PI Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;

DR WPI; 2001-300520/31.

PT New DNA fragments from the rpoB gene of mycobacteria, useful for  
PT diagnosis and identification of many mycobacterial species by  
PT restriction fragment length polymorphism

PS Claim 1; Page 43-44; 50pp; English.

CC The present sequence for Mycobacterium szulgai rpoB gene  
CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from  
CC various Mycobacterial species. These rpoB gene fragments can be used  
CC in the diagnosis and identification of Mycobacterium species using a  
CC novel PCR-restriction fragment length polymorphism analysis (PRA)  
CC method. The method comprises obtaining a restriction fragment length  
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,  
CC amplifying and digesting the DNA fragment from the microorganism to

CC be identified and comparing the RFLP patterns from the known rpoB gene  
CC fragments with the unidentified fragment. The rpoB gene fragments  
CC are useful to identify a wide range of Mycobacterium species, e.g. for  
CC diagnosis or to obtain epidemiological and pathogenesis information for  
CC selection of appropriate therapies, including M. tuberculosis, M. leprae  
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene  
CC fragments is rapid, precise, simple and cost effective (only 1 PCR  
CC required), and can differentiate between many species in a single  
CC experiment, including those difficult to distinguish by usual biochemical  
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for  
CC detecting specific Mycobacterial species.

Sequence 207 BP; 47 A; 63 C; 66 G; 31 T; 0 other;

Query Match 100.0%; Score 19; DB 22; Length 207;  
Best Local Similarity 100.0%; Pred. No. 7.5;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TCACGAGAACGCTACGA 19  
|||||  
Db 1 TCACGAGAACGCTACGA 19

RESULT 4  
AAS05201

ID AAS05201 standard; DNA; 208 BP.

AC AAS05201;

DT 07-SEP-2001 (first entry)

DE Mycobacterium gordonae type I rpoB gene fragment.

KM Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;  
KW PCR-restriction fragment length polymorphism analysis; ds.

OS Mycobacterium gordonae type I.

PN WO200131061-A1.

PD 03-MAY-2001.

PF 27-OCT-2000; 2000WO-KR01223.

PR 27-OCT-1999; 99KR-0046795.

PA (ERUM-) ERUME BIOTECH CO LTD.

PI Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;

DR WPI; 2001-300520/31.

PT New DNA fragments from the rpoB gene of mycobacteria, useful for  
PT diagnosis and identification of many mycobacterial species by  
PT restriction fragment length polymorphism

PS Claim 1; Page 40; 50pp; English.

CC The present sequence for Mycobacterium gordonae type I rpoB gene  
CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from  
CC various Mycobacterial species. These rpoB gene fragments can be used  
CC in the diagnosis and identification of Mycobacterium species using a  
CC novel PCR-restriction fragment length polymorphism analysis (PRA)  
CC method. The method comprises obtaining a restriction fragment length  
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,  
CC amplifying and digesting the DNA fragment from the microorganism to  
CC be identified and comparing the RFLP patterns from the known rpoB gene  
CC fragments with the unidentified fragment. The rpoB gene fragments  
CC are useful to identify a wide range of Mycobacterium species, e.g. for  
CC diagnosis or to obtain epidemiological and pathogenesis information for  
CC selection of appropriate therapies, including M. tuberculosis, M. leprae  
CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected

CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene  
CC fragments is rapid, precise, simple and cost effective (only 1 PCR  
CC required), and can differentiate between many species in a single  
CC experiment, including those difficult to distinguish by usual biochemical  
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for  
CC detecting specific Mycobacterial species.

Sequence 208 BP; 46 A; 71 C; 65 G; 26 T; 0 other;

Query Match 100.0%; Score 19; DB 22; Length 208;  
Best Local Similarity 100.0%; Pred. No. 7.5;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TCACGAGAACGCTACGA 19  
|||||  
Db 1 TCACGAGAACGCTACGA 19

RESULT 5  
AAS05202

ID AAS05202 standard; DNA; 208 BP.

AC AAS05202;

DT 07-SEP-2001 (first entry)

DE Mycobacterium gordonae type II rpoB gene fragment.

KM Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;  
KW PCR-restriction fragment length polymorphism analysis; ds.

OS Mycobacterium gordonae type II.

PN WO200131061-A1.

PD 03-MAY-2001.

PF 27-OCT-2000; 2000WO-KR01223.

PR 27-OCT-1999; 99KR-0046795.

PA (ERUM-) ERUME BIOTECH CO LTD.

PI Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;

DR WPI; 2001-300520/31.

PT New DNA fragments from the rpoB gene of mycobacteria, useful for  
PT diagnosis and identification of many mycobacterial species by  
PT restriction fragment length polymorphism

PS Claim 1; Page 40; 50pp; English.

CC The present sequence for Mycobacterium gordonae type II rpoB gene  
CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from  
CC various Mycobacterial species. These rpoB gene fragments can be used  
CC in the diagnosis and identification of Mycobacterium species using a  
CC novel PCR-restriction fragment length polymorphism analysis (PRA)  
CC method. The method comprises obtaining a restriction fragment length  
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,  
CC amplifying and digesting the DNA fragment from the microorganism to  
CC be identified and comparing the RFLP patterns from the known rpoB gene  
CC fragments with the unidentified fragment. The rpoB gene fragments  
CC are useful to identify a wide range of Mycobacterium species, e.g. for  
CC diagnosis or to obtain epidemiological and pathogenesis information for  
CC selection of appropriate therapies, including M. tuberculosis, M. leprae  
CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected  
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene  
CC fragments is rapid, precise, simple and cost effective (only 1 PCR  
CC required), and can differentiate between many species in a single  
CC experiment, including those difficult to distinguish by usual biochemical  
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for  
CC detecting specific Mycobacterial species.

Query Match	100.0%	Score 19;	DB 22;	Length 208;					
Best Local Similarity	100.0%	Pred. No. 7.5;							
Matches	19;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0

Db  
1 TCAAGGAGAAAGCGCTACGA 19

## RESULT 8

AAS05206 standard; DNA: 208 BP.

AAS05206:

07-SEP-2001 (first entry)

Mycobacterium terrae rpoB gene fragment.

Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP; PCR-restriction fragment length polymorphism analysis; ds.

Mycobacterium terrae.

MO200131061-A1.

03-MAY-2001.

27-OCT-2000; 2000MO-KR01223.

27-OCT-1999; 99KR-0046795.

(ERUM-) ERUME BIOTECH CO LTD.

Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;

WPI: 2001-300520/31.

New DNA fragments from the rpoB gene of mycobacteria, useful for diagnosis and identification of many mycobacterial species by restriction fragment length polymorphism

Claim 1; Page 42; 50pp; English.

The present sequence for Mycobacterium terrae rpoB gene fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from various Mycobacterium species. These rpoB gene fragments can be used in the diagnosis and identification of Mycobacterium species using a novel PCR-restriction fragment length polymorphism analysis (PRA) method. The method comprises obtaining a restriction fragment length polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating, amplifying and digesting the DNA fragment from the microorganism to be identified and comparing the RFLP patterns from the known rpoB gene fragments with the unidentified fragment. The rpoB gene fragments are useful to identify a wide range of Mycobacterium species, e.g. for diagnosis or to obtain epidemiological and pathogenesis information for selection of appropriate therapies, including M. tuberculosis, M. leprae and non-tuberculous mycobacteria (NTM) encountered in subjects infected with human immunodeficiency virus (HIV). Analysis of the rpoB gene fragments is rapid, precise, simple and cost effective (only 1 PCR required), and can differentiate between many species in a single experiment, including those difficult to distinguish by usual biochemical tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for detecting specific Mycobacterium species.

Sequence 208 BP; 46 A; 62 C; 67 G; 33 T; 0 other;

Query Match 100.0%; Score 19; DB 22; Length 208;

Best Local Similarity 100.0%; Pred. No. 7.5; Mismatches 0; Indels 0; Gaps 0;

Matches 19: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TCAGGAGAGCGCTACGA 19  
 |||||||  
 DB 1 TCAGGAGAGCGCTACGA 19

## RESULT 9

AAS05208

AAS05208 standard; DNA: 208 BP.

AAS05208:

XX

07-SEP-2001 (first entry)

Mycobacterium kansasii rpoB gene fragment.

Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP; PCR-restriction fragment length polymorphism analysis; ds.

Mycobacterium kansasii.

MO200131061-A1.

03-MAY-2001.

27-OCT-2000; 2000MO-KR01223.

27-OCT-1999; 99KR-0046795.

(ERUM-) ERUME BIOTECH CO LTD.

Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;

WPI: 2001-300520/31.

New DNA fragments from the rpoB gene of mycobacteria, useful for diagnosis and identification of many mycobacterial species by restriction fragment length polymorphism

Claim 1; Page 42; 50pp; English.

The present sequence for Mycobacterium kansasii rpoB gene fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from various Mycobacterium species. These rpoB gene fragments can be used in the diagnosis and identification of Mycobacterium species using a novel PCR-restriction fragment length polymorphism analysis (PRA) method. The method comprises obtaining a restriction fragment length polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating, amplifying and digesting the DNA fragment from the microorganism to be identified and comparing the RFLP patterns from the known rpoB gene fragments with the unidentified fragment. The rpoB gene fragments are useful to identify a wide range of Mycobacterium species, e.g. for diagnosis or to obtain epidemiological and pathogenesis information for selection of appropriate therapies, including M. tuberculosis, M. leprae and non-tuberculous mycobacteria (NTM) encountered in subjects infected with human immunodeficiency virus (HIV). Analysis of the rpoB gene fragments is rapid, precise, simple and cost effective (only 1 PCR required), and can differentiate between many species in a single experiment, including those difficult to distinguish by usual biochemical tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for detecting specific Mycobacterium species.

Sequence 208 BP; 51 A; 65 C; 65 G; 27 T; 0 other;

Query Match 100.0%; Score 19; DB 22; Length 208;

Best Local Similarity 100.0%; Pred. No. 7.5; Mismatches 0; Indels 0; Gaps 0;

Matches 19: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TCAGGAGAGCGCTACGA 19  
 |||||||  
 DB 1 TCAGGAGAGCGCTACGA 19

## RESULT 10

AAS05210

AAS05210 standard; DNA: 208 BP.

AAS05210:

07-SEP-2001 (first entry)

Mycobacterium ulcerans rpoB gene fragment.

Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP; PCR-restriction fragment length polymorphism analysis; ds.



OS Mycobacterium ulcerans.  
XX  
XX WO200131061-A1.  
XX  
XX 03-MAY-2001.  
XX  
XX 27-OCT-2000; 2000MO-KR01223.  
XX  
XX 27-OCT-1999; 99KR-0046795.  
XX  
XX (ERUM-) ERUME BIOTECH CO LTD.  
XX  
XX Lee H, Park YK, Bal G, Kim S, Cho S, Kim Y, Park HJ;  
XX WPI; 2001-300520/31.  
XX  
XX New DNA fragments from the rpoB gene of mycobacteria, useful for  
XX diagnosis and identification of many mycobacterial species by  
XX restriction fragment length polymorphism -  
XX  
XX Claim 1; Page 43; 50pp; English.  
XX  
XX The present sequence for Mycobacterium ulcerans rpoB gene  
XX fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from  
XX various Mycobacterial species. These rpoB gene fragments can be used  
XX in the diagnosis and identification of Mycobacterium species using a  
XX novel PCR-restriction fragment length polymorphism analysis (PRA)  
XX method. The method comprises obtaining a restriction fragment length  
XX polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,  
XX amplifying and digesting the DNA fragment from the microorganism to  
XX be identified and comparing the RFLP patterns from the known rpoB gene  
XX fragments with the unidentified fragment. The rpoB gene fragments  
XX are useful to identify a wide range of Mycobacterium species, e.g. for  
XX diagnosis or to obtain epidemiological and pathogenesis information for  
XX selection of appropriate therapies, including M. tuberculosis, M. leprae  
XX and non-tuberculous mycobacteria (NTM) encountered in subjects infected  
XX with human immunodeficiency virus (HIV). Analysis of the rpoB gene  
XX fragments is rapid, precise, simple and cost effective (only 1 PCR  
XX required), and can differentiate between many species in a single  
XX experiment, including those difficult to distinguish by usual biochemical  
XX tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for  
XX detecting specific Mycobacterial species.  
XX  
SQ Sequence 208 BP; 45 A; 68 C; 67 G; 28 T; 0 other;  
Query Match 100.0%; Score 19; DB 22; Length 208;  
Best Local Similarity 100.0%; Pred. No. 7.5;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 TCAAGGAGAGCGCTACGA 19  
DB 1 TCAAGGAGAGCGCTACGA 19  
RESULT 11  
AAS05211  
ID AAS05211 standard; DNA: 208 BP.  
XX  
XX AAS05211;  
XX  
XX 07-SEP-2001 (first entry)  
XX  
XX Mycobacterium marinum rpoB gene fragment.  
XX  
XX Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;  
XX PCR-restriction fragment length polymorphism analysis; ds.  
XX  
XX Mycobacterium marinum.  
XX  
XX WO200131061-A1.  
XX  
XX 03-MAY-2001.

XX  
XX 27-OCT-2000; 2000MO-KR01223.  
XX  
XX 27-OCT-1999; 99KR-0046795.  
XX  
XX (ERUM-) ERUME BIOTECH CO LTD.  
XX  
XX Lee H, Park YK, Bal G, Kim S, Cho S, Kim Y, Park HJ;  
XX WPI; 2001-300520/31.  
XX  
XX New DNA fragments from the rpoB gene of mycobacteria, useful for  
XX diagnosis and identification of many mycobacterial species by  
XX restriction fragment length polymorphism -  
XX  
XX Claim 1; Page 43; 50pp; English.

XX The present sequence for Mycobacterium marinum rpoB gene  
XX fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from  
XX various Mycobacterial species. These rpoB gene fragments can be used  
XX in the diagnosis and identification of Mycobacterium species using a  
XX novel PCR-restriction fragment length polymorphism analysis (PRA)  
XX method. The method comprises obtaining a restriction fragment length  
XX polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,  
XX amplifying and digesting the DNA fragment from the microorganism to  
XX be identified and comparing the RFLP patterns from the known rpoB gene  
XX fragments with the unidentified fragment. The rpoB gene fragments  
XX are useful to identify a wide range of Mycobacterium species, e.g. for  
XX diagnosis or to obtain epidemiological and pathogenesis information for  
XX selection of appropriate therapies, including M. tuberculosis, M. leprae  
XX and non-tuberculous mycobacteria (NTM) encountered in subjects infected  
XX with human immunodeficiency virus (HIV). Analysis of the rpoB gene  
XX fragments is rapid, precise, simple and cost effective (only 1 PCR  
XX required), and can differentiate between many species in a single  
XX experiment, including those difficult to distinguish by usual biochemical  
XX tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for  
XX detecting specific Mycobacterial species.

SQ Sequence 208 BP; 45 A; 69 C; 68 G; 26 T; 0 other;

Query Match 100.0%; Score 19; DB 22; Length 208;  
Best Local Similarity 100.0%; Pred. No. 7.5;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TCAAGGAGAGCGCTACGA 19  
DB 1 TCAAGGAGAGCGCTACGA 19

RESULT 12  
AAS05215  
ID AAS05215 standard; DNA: 208 BP.  
XX  
XX AAS05215;  
XX  
XX 07-SEP-2001 (first entry)  
XX  
XX Mycobacterium avium rpoB gene fragment.  
XX  
XX Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;  
XX PCR-restriction fragment length polymorphism analysis; ds.  
XX  
XX Mycobacterium avium.  
XX  
XX WO200131061-A1.  
XX  
XX 03-MAY-2001.  
XX  
XX 27-OCT-2000; 2000MO-KR01223.  
XX  
XX 27-OCT-1999; 99KR-0046795.  
XX  
XX (ERUM-) ERUME BIOTECH CO LTD.

```
XX
PI Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;
XX
DR WPI: 2001-300520/31.
XX
PT New DNA fragments from the rpoB gene of mycobacteria, useful for
PT diagnosis and identification of many mycobacterial species by
PT restriction fragment length polymorphism -
XX
PS Claim 1: Page 44: 50pp: English.
XX
CC The present sequence for Mycobacterium avium rpoB gene
CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
CC various Mycobacterial species. These rpoB gene fragments can be used
CC in the diagnosis and identification of Mycobacterium species using a
CC novel PCR-restriction fragment length polymorphism analysis (PRA)
CC method. The method comprises obtaining a restriction fragment length
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
CC amplifying and digesting the DNA fragment from the microorganism to
CC be identified and comparing the RFLP patterns from the known rpoB gene
CC fragments with the unidentified fragment. The rpoB gene fragments
CC are useful to identify a wide range of Mycobacterium species, e.g. for
CC diagnosis or to obtain epidemiological and pathogenesis information for
CC selection of appropriate therapies, including M. tuberculosis, M. leprae
CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene
CC fragments is rapid, precise, simple and cost effective (only 1 PCR
CC required), and can differentiate between many species in a single
CC experiment, including those difficult to distinguish by usual biochemical
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
CC detecting specific Mycobacterial species.
XX
SQ Sequence 208 BP; 44 A; 69 C; 69 G; 26 T; 0 other;
XX
Query Match 100.0%; Score 19; DB 22; Length 208;
Best Local Similarity 100.0%; Pred. No. 7.5;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 TCACGAGAGAGCGCTACGA 19
DB 1 TCACGAGAGAGCGCTACGA 19
XX
RESULT 13
AAS05216
ID AAS05216 standard; DNA: 208 BP.
XX
AC AAS05216;
XX
DT 07-SEP-2001 (first entry)
XX
DE Mycobacterium bovis rpoB gene fragment.
XX
KM Non-tuberculous mycobacteria: rpoB gene fragment; NTM; HIV; PRA; RFLP;
KM PCR-restriction fragment length polymorphism analysis; ds.
XX
OS Mycobacterium bovis.
XX
PN WO200131061-A1.
XX
PD 03-MAY-2001.
XX
PF 27-OCT-2000; 2000WO-KR01223.
XX
PR 27-OCT-1999; 99KR-0046795.
XX
PA (ERUM-) ERUME BIOTECH CO LTD.
XX
PI Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;
XX
DR WPI: 2001-300520/31.
XX
PT New DNA fragments from the rpoB gene of mycobacteria, useful for
```

```
PT diagnosis and identification of many mycobacterial species by
PT restriction fragment length polymorphism -
XX
PS Claim 1: Page 45; 50pp: English.
XX
CC The present sequence for Mycobacterium bovis rpoB gene
CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
CC various Mycobacterial species. These rpoB gene fragments can be used
CC in the diagnosis and identification of Mycobacterium species using a
CC novel PCR-restriction fragment length polymorphism analysis (PRA)
CC method. The method comprises obtaining a restriction fragment length
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
CC amplifying and digesting the DNA fragment from the microorganism to
CC be identified and comparing the RFLP patterns from the known rpoB gene
CC fragments with the unidentified fragment. The rpoB gene fragments
CC are useful to identify a wide range of Mycobacterium species, e.g. for
CC diagnosis or to obtain epidemiological and pathogenesis information for
CC selection of appropriate therapies, including M. tuberculosis, M. leprae
CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene
CC fragments is rapid, precise, simple and cost effective (only 1 PCR
CC required), and can differentiate between many species in a single
CC experiment, including those difficult to distinguish by usual biochemical
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
CC detecting specific Mycobacterial species.
XX
SQ Sequence 208 BP; 46 A; 62 C; 67 G; 33 T; 0 other;
XX
Query Match 100.0%; Score 19; DB 22; Length 208;
Best Local Similarity 100.0%; Pred. No. 7.5;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 TCACGAGAGAGCGCTACGA 19
DB 1 TCACGAGAGAGCGCTACGA 19
XX
RESULT 14
AAS05217
ID AAS05217 standard; DNA: 208 BP.
XX
AC AAS05217;
XX
DT 07-SEP-2001 (first entry)
XX
DE Mycobacterium celatum rpoB gene fragment.
XX
KM Non-tuberculous mycobacteria: rpoB gene fragment; NTM; HIV; PRA; RFLP;
KM PCR-restriction fragment length polymorphism analysis; ds.
XX
OS Mycobacterium celatum.
XX
PN WO200131061-A1.
XX
PD 03-MAY-2001.
XX
PF 27-OCT-2000; 2000WO-KR01223.
XX
PR 27-OCT-1999; 99KR-0046795.
XX
PA (ERUM-) ERUME BIOTECH CO LTD.
XX
PI Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;
XX
DR WPI: 2001-300520/31.
XX
PT New DNA fragments from the rpoB gene of mycobacteria, useful for
PT diagnosis and identification of many mycobacterial species by
PT restriction fragment length polymorphism -
XX
PS Claim 1: Page 45; 50pp: English.
XX
CC The present sequence for Mycobacterium celatum rpoB gene
```

CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from  
CC various Mycobacterial species. These rpoB gene fragments can be used  
CC in the diagnosis and identification of Mycobacterium species using a  
CC novel PCR-restriction fragment length polymorphism analysis (PRA)  
CC method. The method comprises obtaining a restriction fragment length  
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,  
CC amplifying and digesting the DNA fragment from the microorganism to  
CC be identified and comparing the RFLP patterns from the known rpoB gene  
CC fragments with the unidentified fragment. The rpoB gene fragments  
CC are useful to identify a wide range of Mycobacterium species, e.g. for  
CC diagnosis or to obtain epidemiological and pathogenesis information for  
CC selection of appropriate therapies, including M. tuberculosis, M. leprae  
CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected  
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene  
CC fragments is rapid, precise, simple and cost effective (only 1 PCR  
CC required), and can differentiate between many species in a single  
CC experiment, including those difficult to distinguish by usual biochemical  
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for  
CC detecting specific Mycobacterial species.

XX Sequence 208 BP; 48 A; 71 C; 64 G; 25 T; 0 other;

Query Match 100.0%; Score 19; DB 22; Length 208;

Best Local Similarity 100.0%; Pred. No. 7.5;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TCAAGGAGAAAGCGCTACGA 19

|||||

DB 1 TCAAGGAGAAAGCGCTACGA 19

#### RESULT 15

AAS05219

ID AAS05219 standard: DNA: 208 BP.

AC AAS05219;

XX 07-SEP-2001 (first entry)

DE Mycobacterium fortuitum rpoB gene fragment.

XX Non-tuberculous mycobacteria: rpoB gene fragment; NTM: HIV; PRA; RFLP;

KM PCR-restriction fragment length polymorphism analysis; ds.

XX Mycobacterium fortuitum.

OS WO200131061-A1.

PN 03-MAY-2001.

PD 27-OCT-2000; 2000WO-KR01223.

PR 27-OCT-1999; 99KR-0046795.

PA (ERUM-) ERUME BIOTECH CO LTD.

PI Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;

XX WPI: 2001-300520/31.

PT New DNA fragments from the rpoB gene of mycobacteria, useful for  
PT diagnosis and identification of many Mycobacterial species by  
PT restriction fragment length polymorphism

PS Claim 1; Page 46; 50pp; English.

XX The present sequence for Mycobacterium fortuitum rpoB gene  
CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from  
CC various Mycobacterial species. These rpoB gene fragments can be used  
CC in the diagnosis and identification of Mycobacterium species using a  
CC novel PCR-restriction fragment length polymorphism analysis (PRA)  
CC method. The method comprises obtaining a restriction fragment length  
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,

CC amplifying and digesting the DNA fragment from the microorganism to  
CC be identified and comparing the RFLP patterns from the known rpoB gene  
CC fragments with the unidentified fragment. The rpoB gene fragments  
CC are useful to identify a wide range of Mycobacterium species, e.g. for  
CC diagnosis or to obtain epidemiological and pathogenesis information for  
CC selection of appropriate therapies, including M. tuberculosis, M. leprae  
CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected  
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene  
CC fragments is rapid, precise, simple and cost effective (only 1 PCR  
CC required), and can differentiate between many species in a single  
CC experiment, including those difficult to distinguish by usual biochemical  
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for  
CC detecting specific Mycobacterial species.

XX Sequence 208 BP; 43 A; 69 C; 69 G; 27 T; 0 other;

Query Match 100.0%; Score 19; DB 22; Length 208;

Best Local Similarity 100.0%; Pred. No. 7.5;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TCAAGGAGAAAGCGCTACGA 19

|||||

DB 1 TCAAGGAGAAAGCGCTACGA 19

Search completed: November 12, 2002, 16:50:33  
Job time : 14.4434 secs

GenCore version 5.1.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 12, 2002, 16:50:46 ; Search time 58.0983 Seconds  
(without alignments)  
9517.553 Million cell updates/sec

Title: US-09-697-123b-25

Perfect score: 19  
Sequence: 1 tcaagagaagcgtacga 19

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:\*

1: gb.ba.\*  
2: gb.htg.\*  
3: gb.in.\*  
4: gb.om.\*  
5: gb.ov.\*  
6: gb.pat.\*  
7: gb.ph.\*  
8: gb.pl.\*  
9: gb.pr.\*  
10: gb.ro.\*  
11: gb.sts.\*  
12: gb.sy.\*  
13: gb.un.\*  
14: gb.vl.\*  
15: em.ba.\*  
16: em.fun.\*  
17: em.hum.\*  
18: em.in.\*  
19: em.mu.\*  
20: em.om.\*  
21: em.or.\*  
22: em.ov.\*  
23: em.pat.\*  
24: em.ph.\*  
25: em.pl.\*  
26: em.ro.\*  
27: em.sts.\*  
28: em.un.\*  
29: em.vi.\*  
30: em.htg.hum.\*  
31: em.htg.inv.\*  
32: em.htg.other.\*  
33: em.htg.mus.\*  
34: em.htg.pln.\*  
35: em.htg.rpd.\*  
36: em.htg.mam.\*  
37: em.htg.vrt.\*  
38: em.sy.\*  
39: em.htgo.hum.\*  
40: em.htgo.mus.\*  
41: em.htgo.other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	19	100.0	610	1	MTU318818	AJ318818 Mycobacte
2	19	100.0	610	1	MTU318819	AJ318819 Mycobacte
3	19	100.0	616	1	MTU318813	AJ318813 Mycobacte
4	19	100.0	618	1	MTU318815	AJ318815 Mycobacte
5	19	100.0	618	1	MTU318817	AJ318817 Mycobacte
6	19	100.0	633	1	MTU318814	AJ318814 Mycobacte
7	19	100.0	637	1	MTU318816	AJ318816 Mycobacte
8	19	100.0	639	1	MTU318821	AJ318821 Mycobacte
9	19	100.0	970	6	150706	150706 Sequence 1
10	19	100.0	3534	6	AX111339	AX111339 Sequence
11	19	100.0	3752	1	MSU24494	U24494 Mycobacteri
12	19	100.0	3853	1	MTU12205	U12205 Mycobacteri
13	19	100.0	5084	1	MSGRPOB	L27989 Mycobacteri
14	19	100.0	19352	1	AE006964	AE006964 Mycobacte
15	19	100.0	19770	1	MTU31376	295972 Mycobacteri
16	17.4	91.6	2209	3	AF019754	AF019754 Drosophill
17	17.4	91.6	2238	3	AF057162	AF057162 Drosophill
18	17.4	91.6	2849	3	AF041439	AF041439 Drosophill
19	17.4	91.6	3030	3	AF019753	AF019753 Drosophill
20	17.4	91.6	3244	3	AY069496	AY069496 Drosophill
21	17.4	91.6	3253	3	AF039232	AF039232 Drosophill
22	17.4	91.6	3258	3	AF027729	AF027729 Drosophill
23	17.4	91.6	3447	6	AR067447	AR067447 Sequence
24	17.4	91.6	3941	1	AF242549	AF242549 Amycolato
25	17.4	91.6	4391	3	AF039233	AF039233 Drosophill
26	17.4	91.6	11856	1	AE011788	AE011788 Xanthomon
27	17.4	91.6	37617	1	MLB1790G	214314 M.leprae ge
28	17.4	91.6	52259	2	AC017309	AC017309 Drosophill
29	17.4	91.6	124039	2	AC019900	AC019900 Drosophill
30	17.4	91.6	154298	8	AP003504	AP003504 Oryza sat
31	17.4	91.6	159528	3	AC008310	AC008310 Drosophill
32	17.4	91.6	174036	3	AC010214	AC010214 Drosophill
33	17.4	91.6	232064	3	AE003779	AE003779 Drosophill
34	17.4	91.6	253648	3	AE003638	AE003638 Drosophill
35	17.4	91.6	348950	1	MLEPRTN7	AL583923 Mycobacte
36	16.4	86.3	2045	5	FSADB	D55629 Sea lamprey
37	16.4	86.3	3691	9	BC026311	BC026311 Homo sapi
38	16.4	86.3	5647	2	AC014154	AC014154 Drosophill
39	16.4	86.3	10306	1	AE005712	AE005712 Caulobact
40	16.4	86.3	10393	1	AE005136	AE005136 Halobacte
41	16.4	86.3	11021	1	AE004674	AE004674 Pseudomon
42	16.4	86.3	36642	1	AF029344	AF029344 Burkholder
43	16.4	86.3	39896	1	SKK13	AL512667 Streptomy
44	16.4	86.3	41399	2	AC112021	AC112021 Rattus no
45	16.4	86.3	42210	1	SC1C2	AL031124 Streptomy

# ALIGNMENTS

RESULT 1  
LOCUS MTU318818  
DEFINITION Mycobacterium tuberculosis partial rpoB gene for RNA polymerase  
beta subunit, isolate 1415-97.  
ACCESSION AJ318818  
VERSION AJ318818.1 GI:22208412  
KEYWORDS RNA polymerase beta subunit; rpoB gene.  
SOURCE Mycobacterium tuberculosis.  
ORGANISM Mycobacterium tuberculosis  
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium  
tuberculosis complex.  
REFERENCE 1  
AUTHORS Herrera, L., Jimenez, M.S. and Saez, J.A.

MTU318818 610 bp DNA circular BCT 09-AUG-2002  
Mycobacterium tuberculosis partial rpoB gene for RNA polymerase  
beta subunit, isolate 1415-97.

```

TITLE
Molecular analysis of rifampin-resistant Mycobacterium tuberculosis
isolated in Spain (1996-2001). Description of new alleles into rpoB
gene and review
JOURNAL
REFERENCE
2 (bases 1 to 610)
Unpublished
AUTHORS
Herrera,L.
TITLE
Direct Submission
JOURNAL
Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro
Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5,
Majadahonda. Madrid. 28220, SPAIN
FEATURES
source
1..610
Location/Qualifiers
/organism="Mycobacterium tuberculosis"
/isolate="1415-97"
/db_xref="taxon:1773"
gene
1..610
/genes="rpoB"
<1..>610
CDS
<1..>610
/codon_start=1
/transl_table=11
/product="RNA polymerase beta subunit"
/protein_id="CAC87035.1"
/db_xref="GI:22208413"
/translation="LDIYRKLRPGEPTKESQTLLENLFFKERYDLARVGRYKVNK
KLGHWGEPTSTLTEDVVAATIEYVRLHGGOTTMTVPGGVVETDDIDHFGNR
RLRTVGELIQNIRVGMSEVRVRRMTQDVEAITPOTLINIRPVVAATKEFFGTS
QPSQFMGQNNPLSLGTHKRRLSALGPGLSRERAGLEVRDVHP"
BASE COUNT 122 a 191 c 202 g 95 t
ORIGIN
Query Match 100.0%; Score 19; DB 1; Length 610;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TCAAGGAGAGCGCTACGA 19
|||||
Db 80 TCAAGGAGAGCGCTACGA 98
|||||

RESULT 2
MTU318819
LOCUS
Mycobacterium tuberculosis partial rpoB gene for RNA polymerase
beta subunit, isolate 1417-97.
DEFINITION
ACCESSION
AJ318819
VERSION
1 GI:22208414
KEYWORDS
RNA polymerase beta subunit; rpoB gene.
SOURCE
Mycobacterium tuberculosis.
ORGANISM
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
tuberculosis complex.
REFERENCE
1
AUTHORS
Herrera,L., Jimenez,M.S. and Saez,J.A.
TITLE
Molecular analysis of rifampin-resistant Mycobacterium tuberculosis
isolated in Spain (1996-2001). Description of new alleles into rpoB
gene and review
JOURNAL
REFERENCE
2 (bases 1 to 610)
AUTHORS
Herrera,L.
TITLE
Direct Submission
JOURNAL
Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro
Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5,
Majadahonda. Madrid. 28220, SPAIN
FEATURES
source
1..610
Location/Qualifiers
/organism="Mycobacterium tuberculosis"
/isolate="1417-97"
/db_xref="taxon:1773"
gene
1..610
/genes="rpoB"
<1..>610
CDS
<1..>610

```

```

/genes="rpoB"
/codon_start=1
/transl_table=11
/product="RNA polymerase beta subunit"
/protein_id="CAC87036.1"
/db_xref="GI:22208415"
/translation="LDIYRKLRPGEPTKESQTLLENLFFKERYDLARVGRYKVNK
KLGHWGEPTSTLTEDVVAATIEYVRLHGGOTTMTVPGGVVETDDIDHFGNR
RLRTVGELIQNIRVGMSEVRVRRMTQDVEAITPOTLINIRPVVAATKEFFGTS
QPSQFMGQNNPLSLGTHKRRLSALGPGLSRERAGLEVRDVHP"
BASE COUNT 122 a 191 c 202 g 95 t
ORIGIN
Query Match 100.0%; Score 19; DB 1; Length 610;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TCAAGGAGAGCGCTACGA 19
|||||
Db 80 TCAAGGAGAGCGCTACGA 98
|||||

RESULT 3
MTU318813
LOCUS
Mycobacterium tuberculosis partial rpoB gene for RNA polymerase
beta subunit, isolate 1763-97.
DEFINITION
ACCESSION
AJ318813
VERSION
1 GI:22208402
KEYWORDS
RNA polymerase beta subunit; rpoB gene.
SOURCE
Mycobacterium tuberculosis.
ORGANISM
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
tuberculosis complex.
REFERENCE
1
AUTHORS
Herrera,L., Jimenez,M.S. and Saez,J.A.
TITLE
Molecular analysis of rifampin-resistant Mycobacterium tuberculosis
isolated in Spain (1996-2001). Description of new alleles into rpoB
gene and review
JOURNAL
REFERENCE
2 (bases 1 to 616)
AUTHORS
Herrera,L.
TITLE
Direct Submission
JOURNAL
Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro
Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5,
Majadahonda. Madrid. 28220, SPAIN
FEATURES
source
1..616
Location/Qualifiers
/organism="Mycobacterium tuberculosis"
/isolate="1763-97"
/db_xref="taxon:1773"
gene
1..615
<1..>615
CDS
<1..>615
/genes="rpoB"
/codon_start=1
/transl_table=11
/product="RNA polymerase beta subunit"
/protein_id="CAC87030.1"
/db_xref="GI:22208403"
/translation="LDIYRKLRPGEPTKESQTLLENLFFKERYDLARVGRYKVNK
KLGHWGEPTSTLTEDVVAATIEYVRLHGGOTTMTVPGGVVETDDIDHFGNR
RLRTVGELIQNIRVGMSEVRVRRMTQDVEAITPOTLINIRPVVAATKEFFGTS
QLSQIFMDQNNPLSLGTHKRRLSALGPGLSRERAGLEVRDVHP"
BASE COUNT 125 a 191 c 201 g 99 t
ORIGIN
Query Match 100.0%; Score 19; DB 1; Length 616;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TCAAGGAGAGCGCTACGA 19

```

```

|||||
Db 80 TCAAGGAGAGCGCTACGA 98

RESULT 4
MTU318815
LOCUS
DEFINITION
Mycobacterium tuberculosis partial rpoB gene for RNA polymerase
beta subunit, isolate 2540-97.
ACCESSION
AJ318815
VERSION
AJ318815.1 GI:22208406
KEYWORDS
RNA polymerase beta subunit; rpoB gene.
SOURCE
Mycobacterium tuberculosis.
ORGANISM
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
tuberculosis complex.
REFERENCE
1
AUTHORS
Herrera,L., Jimenez,M.S. and Saez,J.A.
TITLE
Molecular analysis of rifampin-resistant Mycobacterium tuberculosis
isolated in Spain (1996-2001). Description of new alleles into rpoB
gene and review
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 618)
AUTHORS
Herrera,L.
TITLE
Direct Submission
JOURNAL
Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro
Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5,
Majadahonda. Madrid. 28220, SPAIN
FEATURES
source
1..618
/organism="Mycobacterium tuberculosis"
/isolate="2540-97"
/db_xref="taxon:1773"
gene
1..618
/gene="rpoB"
CDS
<1..>618
/gene="rpoB"
/codon_start=1
/transl_table=11
/product="RNA polymerase beta subunit"
/protein_id="CAC87032.1"
/db_xref="GI:22208407"
/translation="TDEALLDIYRKLPGPEPTKESAOITLLENLFFKRYDLARVGR
YKVNKGLGHVEPITSTLDEEVATIEYLVRLHGGQTTMTVPGVGVPEVETDDID
HFGNRLRTVGLIQNIQVGMSEVVRMTQDVEAITPRLINIRPVAAIKE
FFGTSQSFQMDQNNPLSGLTTHKRLLALGPGLSRERAGLEVDRV"
BASE COUNT 124 a 190 c 207 g 97 t
ORIGIN
/organism="Mycobacterium tuberculosis"
/isolate="2540-97"
/db_xref="taxon:1773"
1..618
/gene="rpoB"
<1..>618
/gene="rpoB"
/codon_start=1
/transl_table=11
/product="RNA polymerase beta subunit"
/protein_id="CAC87032.1"
/db_xref="GI:22208407"
/translation="TDEALLDIYRKLPGPEPTKESAOITLLENLFFKRYDLARVGR
YKVNKGLGHVEPITSTLDEEVATIEYLVRLHGGQTTMTVPGVGVPEVETDDID
HFGNRLRTVGLIQNIQVGMSEVVRMTQDVEAITPRLINIRPVAAIKE
FFGTSQSFQMDQNNPLSGLTTHKRLLALGPGLSRERAGLEVDRV"
BASE COUNT 124 a 190 c 207 g 97 t
ORIGIN
Query Match 100.0%; Score 19; DB 1; Length 618;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TCAAGGAGAGCGCTACGA 19
|||||
Db 95 TCAAGGAGAGCGCTACGA 113

RESULT 5
MTU318817
LOCUS
DEFINITION
Mycobacterium tuberculosis partial rpoB gene for RNA polymerase
beta subunit, isolate 2348-98.
ACCESSION
AJ318817
VERSION
AJ318817.1 GI:22208410
KEYWORDS
RNA polymerase beta subunit; rpoB gene.
SOURCE
Mycobacterium tuberculosis.
ORGANISM
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
tuberculosis complex.
REFERENCE
1
AUTHORS
Herrera,L., Jimenez,M.S. and Saez,J.A.
TITLE
Molecular analysis of rifampin-resistant Mycobacterium tuberculosis
isolated in Spain (1996-2001). Description of new alleles into rpoB
gene and review
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 633)
AUTHORS
Herrera,L.
TITLE
Direct Submission
JOURNAL
Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro
Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5,
Majadahonda. Madrid. 28220, SPAIN
FEATURES
source
1..633
/organism="Mycobacterium tuberculosis"
/isolate="1058-97"
/db_xref="taxon:1773"
1..633
/gene="rpoB"
gene
1..633
/gene="rpoB"
CDS
<1..>633
/gene="rpoB"
/codon_start=1
/transl_table=11
/product="RNA polymerase beta subunit"
/protein_id="CAC87034.1"
/db_xref="GI:22208411"
/translation="TDEALLDIYRKLPGPEPTKESAOITLLENLFFKRYDLARVGR
YKVNKGLGHVEPITSTLDEEVATIEYLVRLHGGQTTMTVPGVGVPEVETDDID
HFGNRLRTVGLIQNIQVGMSEVVRMTQDVEAITPRLINIRPVAAIKE
FFGTSQSFQMDQNNPLSGLTTHKRLLALGPGLSRERAGLEVDRV"
BASE COUNT 124 a 192 c 207 g 95 t
ORIGIN
/organism="Mycobacterium tuberculosis"
/isolate="2348-98"
/db_xref="taxon:1773"
1..618
/gene="rpoB"
<1..>618
/gene="rpoB"
/codon_start=1
/transl_table=11
/product="RNA polymerase beta subunit"
/protein_id="CAC87034.1"
/db_xref="GI:22208411"
/translation="TDEALLDIYRKLPGPEPTKESAOITLLENLFFKRYDLARVGR
YKVNKGLGHVEPITSTLDEEVATIEYLVRLHGGQTTMTVPGVGVPEVETDDID
HFGNRLRTVGLIQNIQVGMSEVVRMTQDVEAITPRLINIRPVAAIKE
FFGTSQSFQMDQNNPLSGLTTHKRLLALGPGLSRERAGLEVDRV"
BASE COUNT 124 a 192 c 207 g 95 t
ORIGIN
Query Match 100.0%; Score 19; DB 1; Length 618;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TCAAGGAGAGCGCTACGA 19
|||||
Db 95 TCAAGGAGAGCGCTACGA 113

RESULT 6
MTU318814
LOCUS
DEFINITION
Mycobacterium tuberculosis partial rpoB gene for RNA polymerase
beta subunit, isolate 1058-97.
ACCESSION
AJ318814
VERSION
AJ318814.1 GI:22208404
KEYWORDS
RNA polymerase beta subunit; rpoB gene.
SOURCE
Mycobacterium tuberculosis.
ORGANISM
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
tuberculosis complex.
REFERENCE
1
AUTHORS
Herrera,L., Jimenez,M.S. and Saez,J.A.
TITLE
Molecular analysis of rifampin-resistant Mycobacterium tuberculosis
isolated in Spain (1996-2001). Description of new alleles into rpoB
gene and review
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 633)
AUTHORS
Herrera,L.
TITLE
Direct Submission
JOURNAL
Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro
Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5,
Majadahonda. Madrid. 28220, SPAIN
FEATURES
source
1..633
/organism="Mycobacterium tuberculosis"
/isolate="1058-97"
/db_xref="taxon:1773"
1..633
/gene="rpoB"
gene
1..633
/gene="rpoB"
CDS
<1..>633
/gene="rpoB"
/codon_start=1
/transl_table=11
/product="RNA polymerase beta subunit"
/protein_id="CAC87034.1"
/db_xref="GI:22208411"
/translation="TDEALLDIYRKLPGPEPTKESAOITLLENLFFKRYDLARVGR
YKVNKGLGHVEPITSTLDEEVATIEYLVRLHGGQTTMTVPGVGVPEVETDDID
HFGNRLRTVGLIQNIQVGMSEVVRMTQDVEAITPRLINIRPVAAIKE
FFGTSQSFQMDQNNPLSGLTTHKRLLALGPGLSRERAGLEVDRV"
BASE COUNT 124 a 192 c 207 g 95 t
ORIGIN
/organism="Mycobacterium tuberculosis"
/isolate="2348-98"
/db_xref="taxon:1773"
1..618
/gene="rpoB"
<1..>618
/gene="rpoB"
/codon_start=1
/transl_table=11
/product="RNA polymerase beta subunit"
/protein_id="CAC87034.1"
/db_xref="GI:22208411"
/translation="TDEALLDIYRKLPGPEPTKESAOITLLENLFFKRYDLARVGR
YKVNKGLGHVEPITSTLDEEVATIEYLVRLHGGQTTMTVPGVGVPEVETDDID
HFGNRLRTVGLIQNIQVGMSEVVRMTQDVEAITPRLINIRPVAAIKE
FFGTSQSFQMDQNNPLSGLTTHKRLLALGPGLSRERAGLEVDRV"
BASE COUNT 124 a 192 c 207 g 95 t
ORIGIN
Query Match 100.0%; Score 19; DB 1; Length 618;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TCAAGGAGAGCGCTACGA 19
|||||
Db 95 TCAAGGAGAGCGCTACGA 113

RESULT 7
MTU318817
LOCUS
DEFINITION
Mycobacterium tuberculosis partial rpoB gene for RNA polymerase
beta subunit, isolate 2348-98.
ACCESSION
AJ318817
VERSION
AJ318817.1 GI:22208410
KEYWORDS
RNA polymerase beta subunit; rpoB gene.
SOURCE
Mycobacterium tuberculosis.
ORGANISM
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
tuberculosis complex.
REFERENCE
1
AUTHORS
Herrera,L., Jimenez,M.S. and Saez,J.A.
TITLE
Molecular analysis of rifampin-resistant Mycobacterium tuberculosis
isolated in Spain (1996-2001). Description of new alleles into rpoB
gene and review
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 633)
AUTHORS
Herrera,L.
TITLE
Direct Submission
JOURNAL
Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro
Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5,
Majadahonda. Madrid. 28220, SPAIN
FEATURES
source
1..633
/organism="Mycobacterium tuberculosis"
/isolate="1058-97"
/db_xref="taxon:1773"
1..633
/gene="rpoB"
gene
1..633
/gene="rpoB"
CDS
<1..>633
/gene="rpoB"
/codon_start=1
/transl_table=11
/product="RNA polymerase beta subunit"
/protein_id="CAC87034.1"
/db_xref="GI:22208411"
/translation="TDEALLDIYRKLPGPEPTKESAOITLLENLFFKRYDLARVGR
YKVNKGLGHVEPITSTLDEEVATIEYLVRLHGGQTTMTVPGVGVPEVETDDID
HFGNRLRTVGLIQNIQVGMSEVVRMTQDVEAITPRLINIRPVAAIKE
FFGTSQSFQMDQNNPLSGLTTHKRLLALGPGLSRERAGLEVDRV"
BASE COUNT 124 a 192 c 207 g 95 t
ORIGIN
/organism="Mycobacterium tuberculosis"
/isolate="2348-98"
/db_xref="taxon:1773"
1..618
/gene="rpoB"
<1..>618
/gene="rpoB"
/codon_start=1
/transl_table=11
/product="RNA polymerase beta subunit"
/protein_id="CAC87034.1"
/db_xref="GI:22208411"
/translation="TDEALLDIYRKLPGPEPTKESAOITLLENLFFKRYDLARVGR
YKVNKGLGHVEPITSTLDEEVATIEYLVRLHGGQTTMTVPGVGVPEVETDDID
HFGNRLRTVGLIQNIQVGMSEVVRMTQDVEAITPRLINIRPVAAIKE
FFGTSQSFQMDQNNPLSGLTTHKRLLALGPGLSRERAGLEVDRV"
BASE COUNT 124 a 192 c 207 g 95 t
ORIGIN
Query Match 100.0%; Score 19; DB 1; Length 618;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TCAAGGAGAGCGCTACGA 19
|||||
Db 95 TCAAGGAGAGCGCTACGA 113

```

```

CDS
<1..>633
/gene="rpoB"
/codon_start=1
/transl_table=11
/product="RNA polymerase beta subunit"
/protein_id="CAC87031.1"
/db_xref="GI:22208405"
/translation="DNTVGDEALLDIYKLRPGESPPTKESAOTLLENLFFKRYDLARV
ARGRYKVKKGLGHGEPITSTLTEDVVAITYLVRLHGGOTTMTVPGGVPEVDE
TDDIHFGNRLRTVGLIQNIQIRVGMSRMERVVRMTTQDVEAITPQTILNIRPVV
AAIKFEFTSOLSQFMDQNNPLSGLYKRLSALGPGGLSRERAGLEVRDV"
BASE COUNT      129 a 195 c 210 g 99 t
ORIGIN

Query Match      100.0%; Score 19; DB 1; Length 633;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 TCAAGGAGAGCGCTACGA 19
|||||
Db 110 TCAAGGAGAGCGCTACGA 128

RESULT 7
MTU318816
LOCUS
DEFINITION
Mycobacterium tuberculosis partial rpoB gene for RNA polymerase
beta subunit, isolate 1255-98.
ACCESSION
AJ318816
VERSION
AJ318816.1 GI:22208408
KEYWORDS
RNA polymerase beta subunit; rpoB gene.
SOURCE
Mycobacterium tuberculosis.
ORGANISM
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
tuberculosis complex.
REFERENCE
1 Herrera,L., Jimenez,M.S. and Saez,J.A.
Molecular analysis of rifampin-resistant Mycobacterium tuberculosis
isolated in Spain (1996-2001). Description of new alleles into rpoB
gene and review
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 637)
AUTHORS
Herrera,L.
TITLE
Direct Submission
JOURNAL
Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro
Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5,
Majadahonda. Madrid. 28220, SPAIN
FEATURES
source
1..637
/organism="Mycobacterium tuberculosis"
/isolate="1255-98"
/db_xref="taxon:1773"
/gene="rpoB"
CDS
<1..>637
/gene="rpoB"
/codon_start=1
/transl_table=11
/product="RNA polymerase beta subunit"
/protein_id="CAC87031.1"
/db_xref="GI:22208409"
/translation="VGTDEALLDIYKLRPGEPPTKESAOTLLENLFFKRYDLARV
GRYKVKKGLGHGEPITSTLTEDVVAITYLVRLHGGOTTMTVPGGVPEVDE
IDHFGNRLRTVGLIQNIQIRVGMSRMERVVRMTTQDVEAITPQTILNIRPVVAAI
KEFTGSOLSQFMDQNNPLSGLYKRLSALGPGGLSRERAGLEVRDVHPSH"
BASE COUNT      128 a 198 c 210 g 101 t
ORIGIN

Query Match      100.0%; Score 19; DB 1; Length 637;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 TCAAGGAGAGCGCTACGA 19
|||||
Db 101 TCAAGGAGAGCGCTACGA 119

RESULT 8
MTU318821
LOCUS
DEFINITION
Mycobacterium tuberculosis partial rpoB gene for RNA polymerase
beta subunit, isolate 1071-98.
ACCESSION
AJ318821
VERSION
AJ318821.1 GI:22208418
KEYWORDS
RNA polymerase beta subunit; rpoB gene.
SOURCE
Mycobacterium tuberculosis.
ORGANISM
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
tuberculosis complex.
REFERENCE
1 Herrera,L., Jimenez,M.S. and Saez,J.A.
Molecular analysis of rifampin-resistant Mycobacterium tuberculosis
isolated in Spain (1996-2001). Description of new alleles into rpoB
gene and review
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 639)
AUTHORS
Herrera,L.
TITLE
Direct Submission
JOURNAL
Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro
Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5,
Majadahonda. Madrid. 28220, SPAIN
FEATURES
source
1..639
/organism="Mycobacterium tuberculosis"
/isolate="1071-98"
/db_xref="taxon:1773"
/gene="rpoB"
CDS
<1..>639
/gene="rpoB"
/codon_start=1
/transl_table=11
/product="RNA polymerase beta subunit"
/protein_id="CAC87038.1"
/db_xref="GI:22208419"
/translation="RRTDEALLDIYKLRPGEPPTKESAOTLLENLFFKRYDLARV
GRYKVKKGLGHGEPITSTLTEDVVAITYLVRLHGGOTTMTVPGGVPEVDE
IDHFGNRLRTVGLIQNIQIRVGMSRMERVVRMTTQDVEAITPQTILNIRPVVAAI
KEFTGSOLSQFMDQNNPLSGLYKRLSALGPGGLSRERAGLEVRDVHPSH"
BASE COUNT      126 a 202 c 212 g 99 t
ORIGIN

Query Match      100.0%; Score 19; DB 1; Length 639;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 TCAAGGAGAGCGCTACGA 19
|||||
Db 101 TCAAGGAGAGCGCTACGA 119

RESULT 9
I50706
LOCUS
DEFINITION
Sequence 1 from patent US 5643723.
ACCESSION
I50706
VERSION
I50706.1 GI:2472409
KEYWORDS
.
SOURCE
Unknown.
ORGANISM
Unclassified.
REFERENCE
1 (bases 1 to 970)
AUTHORS
Persing,D.H., Hunt,J.J., Young,K.K.Y., Feilmelee,T.A., Roberts,G.D.
and Whelan,A.Christian.

```

TITLE Detection of a genetic locus encoding resistance to rifampin in mycobacterial cultures and in clinical specimens  
JOURNAL Patent: US 5643723-A 1 01-JUL-1997;  
FEATURES Location/Qualifiers  
source 1..970  
/organism="unknown"  
BASE COUNT 182 a 302 c 330 g 156 t  
ORIGIN

Query Match 100.0%; Score 19; DB 6; Length 970;  
Best Local Similarity 100.0%; Pred. No. 36;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCAAGGAGAGCGCTACGA 19  
|||||  
Db 26 TCAAGGAGAGCGCTACGA 44

RESULT 10  
AX111339  
LOCUS  
DEFINITION Sequence 2072 from Patent WO0123604.  
ACCESSION AX111339  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Mycobacterium tuberculosis.  
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
Actinomycetales; Corynebacterineae; Mycobacteriaceae;  
Mycobacterium; Mycobacterium tuberculosis complex.  
1 (bases 1 to 3534)  
Bergerson, M.G., Boissinot, M., Huletsky, A., m Nard, C., Ouellette, M.,  
Picard, F.J. and Roy, P.H.  
Highly conserved genes and their use to generate probes and primers  
for detection of microorganisms  
Patent: WO 0123604-A 2072 05-APR-2001;  
Infectio Diagnostic (I.D.I.) INC. (CA)  
Location/Qualifiers  
source 1..3534  
/organism="Mycobacterium tuberculosis"  
/strain="Rv"  
/db\_xref="taxon:1773"  
BASE COUNT 679 a 1081 c 1188 g 586 t  
ORIGIN

Query Match 100.0%; Score 19; DB 6; Length 3534;  
Best Local Similarity 100.0%; Pred. No. 36;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCAAGGAGAGCGCTACGA 19  
|||||  
Db 902 TCAAGGAGAGCGCTACGA 920

RESULT 11  
MSU24494  
LOCUS  
DEFINITION 3752 bp DNA linear BCT 02-MAR-2000  
ACCESSION U24494  
VERSION U24494.1 GI:790347  
KEYWORDS  
SOURCE  
ORGANISM  
Mycobacterium smegmatis.  
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
Actinomycetales; Corynebacterineae; Mycobacteriaceae;  
Mycobacterium.  
1 (bases 1 to 3752)  
Hetherington, S.V., Watson, A.S. and Patrick, C.C.  
Sequence and analysis of the rpoB gene of Mycobacterium smegmatis  
Antimicrob. Agents Chemother. 39 (9), 2164-2166 (1995)  
MEDLINE 96050766  
PUBMED 8540740  
REFERENCE 2 (bases 1 to 3752)

AUTHORS Hetherington, S.V.  
TITLE Direct Submission  
JOURNAL Submitted (11-APR-1995) Seth V. Hetherington, Infectious Diseases,  
St. Jude Children's Research Hospital, 332 N. Lauderdale, Memphis,  
TN 38101, USA  
FEATURES Location/Qualifiers  
source 1..3752  
/organism="Mycobacterium smegmatis"  
/db\_xref="taxon:1772"  
194..3703  
/gene="rpoB"  
194..3703  
/gene="rpoB"  
/codon\_start=1  
/evidence=experimental  
/transl\_table=11  
/product="DNA polymerase"  
/protein\_id="AAA91426.2"  
/db\_xref="GI:7144498"

LDVOTDSFEMLVGSRVRRQAIDRGEENPVGGLEEVLAELSPIDEGSGMSLSFSDPR  
FDEVKASVDECKDKDWTYAAPLVTAETINNTEIGIKSQTFVGMDFPMMTEKGTFIIN  
GTERVVSQVLRSPGVTFDETDKTSKTEKTLHSVKVIPGRGAWLEFDDVKRDTVGVRIID  
RKRQPVTVLLKALGMTNEQIVERFGSEIMMGTLEKDTTSGTDEALLDIYKLRPGE  
YLVRLHREGQTSMTVPGGVFAVESHDIHFGRNRLRTVGLIONQIRVLSRMERVVR  
ERMTTODVEALITPQTLINIRPVVAAIKEFFGTSQLSQFMDQNNPLSGLTHKRLSLYR  
PGGLSRERAGLEVRDHPHSYHGMCFIETPEGNIGLIGLSLYARVNPFGFTETPYR  
KYVGVSDQIDLYLTADEEDRHVVAQANSPSTDENGRTEDRVNVRKKKGVEFEVSADQ  
VDMVSPRQMVSVATAMIPFLHDDANRALMGANRQADRVPLVRSEAPLVGTGMELR  
AAIDAVATSSHKTGTVIEVSADYITVMADGTQSRVLRKRFARSNHGTSCANORPIVDA  
GORVEAGVIADGCTQNGEMALCNLLVAIMPWEHNYEDAILLSNRLVEEDVLTISI  
HIEHEITDEARDTKLGAETITRVIPKLSDEVLDLDSRIVRIGAEVRDSEILYGVKVP  
KGTETLPEERLLRAGFEKAREVDTSLKVPHGSEKVGIVRFSREDDDELUPGVGN  
ELRVVYAAQRKISDGDKLARRHNGKVIKILPVEDMPFDPDGTVDILNTHGCRV  
VLNQILLETHLGWAGAGNIDVLGAVPDWASKLPPELYSAPADSTVATVPFDGAE  
GELAGLSTLNRDGEVMVNAADKATLDFGRSSGPPFPYPTVGVYIYLKHLHVD  
KIHARSTGYSMITQPLGKAGQFGGGRFCEMCWMAQYGAAYTLQELLTIKSDDTV  
GRVYVEAIVGNIPEPGIPESFKVLLKELQSLCLNVEVLSSDGRAIEMRGGDDDEL  
ERAAANLGINLSRNESASVEDLA"  
BASE COUNT 706 a 1225 c 1210 g 611 t  
ORIGIN

Query Match 100.0%; Score 19; DB 1; Length 3752;  
Best Local Similarity 100.0%; Pred. No. 36;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCAAGGAGAGCGCTACGA 19  
|||||  
Db 1068 TCAAGGAGAGCGCTACGA 1086

RESULT 12  
MTU12205  
LOCUS  
DEFINITION 3853 bp DNA linear BCT 02-MAR-2000  
ACCESSION MTU12205  
VERSION U12205  
KEYWORDS  
SOURCE  
ORGANISM  
Mycobacterium tuberculosis.  
Mycobacterium tuberculosis  
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
Actinomycetales; Corynebacterineae; Mycobacteriaceae;  
Mycobacterium; Mycobacterium tuberculosis complex.  
1 (bases 1 to 3853)  
Imboden, P., Troller, R., Marchesi, F., Telenti, A., Bodmer, T.,  
Cole, S., Schopfer, K. and Burkart, T.  
The rpoB gene of Mycobacterium tuberculosis  
Unpublished  
REFERENCE 2 (bases 1 to 3853)  
Imboden, P.  
Direct Submission



JOURNAL Submitted (11-JUL-1994) Paul Imboden, Institute for Medical Microbiology, University of Berne, Friedbuehlstrasse 51, Berne, 3010, Switzerland

FEATURES  
source 1. 3853  
Location/Qualifiers  
/organism="Mycobacterium tuberculosis"  
/strain="H37Rv"  
/db\_xref="taxon:1773"  
576 .>3853  
/gene="rpoB"  
576 .>3853  
/gene="rpoB"  
/codon\_start=1  
/transl\_table=1  
/product="RNA polymerase beta subunit"  
/protein\_id="AA20242.2"  
/db\_xref="GI:7144499"  
/translation="MLEGCILADSRQSKTAASPSRPSQSSNNNSVPGAPNRVSFAKL  
REPLEVGLLDVDTGSFELIGSPWRRESAAERGVNPGVGLLEVLVLSPIEDFSGS  
MSLSFSDPRFDVVKAPVDECKDKDMYTAAPLFTVAETNNNTGKTSQVPMGDFPMH  
TEKGTFLINGTERVVSQVSPGVVFDETIDKSDTKLHSVKVIPSRGAWLEFDVVK  
RDTGVGRIDRKRQPVTVLLKALGHTSQIIVERGFSFSEIMRSTLEKONTVGTDEALLD  
IYKLRPGEPPTKESAOQLLENLFKERRDLARVGRYKNNKLGHLHVGEPITTSSTLT  
EDDVATIEYLVRHLEGQTTMTVPGVEVPTETDDIDHFNRRLLRTVGEIQNGIRVG  
MSRMERVRERTTQDVEAITPQTLINIRPVYAAIKETFEFTSOLFQMDQNNPLSGIT  
HKRRSALGPGLSRERAGLEVDRVHPHYGRMCPKETPEGNIGLIGSLSVYARVNP  
FGFIETPYKVVGVSDGIVVLTADSDHVAQANSPIDADGRFVEPRVLRKAG  
EVEYVPSSEVDYMDSPROMYSVATAMIPLEHDDANRALGMANKQARQVPLVSRSEAP  
LVGTGMLRAAIDAATSSQESGVIEEVSADYIIVMDHNGTRRTYRMRKFARNHGTCP  
ANQCPVLDAGDRVEAGQVIADGCTDDGEMALGNLLVAIMPWEGHNYEDAIIILSNRL  
VEDVLTSHIEHEIDARDTKLGAETTRDIPNISDEVLDLDERGIVRIGAEVRDG  
DILGKVTGPKETELTPEERLLRAIFGKAREVRDTSIKVPHGESGKVGIRVFSRED  
EDELPAQVNELVYVVAQKRKISDGLAGRHNGKVGIGKILPVEDMPFLADGTPVDI  
ILNTHGVPRRNIGQILETHLGCWAHSGWKVDAAGVPMNAARLPDELLLEAQAIVS  
TPVFGCAEAELOGLSCTLPNRDGDVLVDAGKAMLPDGRSGEPPFPYVTVGWTIM  
KLHLVDDKIHARSTGPGYSMITQPLGKAGQGGQRFEGMECMAQVAGAYTLQELL  
TIKS"

BASE COUNT 723 a 1173 c 1293 g 664 t

ORIGIN

Query Match 100.0%; Score 19; DB 1; Length 3853;  
Best Local Similarity 100.0%; Pred. No. 36;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCAAGGAGAGCGCTACGA 19  
|||||

Db 1477 TCAAGGAGAGCGCTACGA 1495

RESULT 13  
MSGRPOB 5084 bp DNA linear BCT 13-SEP-1994  
LOCUS Mycobacterium tuberculosis RNA polymerase beta-subunit (rpoB)  
DEFINITION gene, complete cds and RNA polymerase beta'-subunit rpoC gene,  
partial cds.  
L27989  
ACCESSION L27989.1 GI:468333  
VERSION RNA polymerase beta-subunit; rpoB gene.  
KEYWORDS Mycobacterium tuberculosis (strain Rv) DNA.  
SOURCE Mycobacterium tuberculosis  
ORGANISM Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium  
tuberculosis complex.  
1 (bases 1 to 5084)  
REFERENCE Miller,L.P., Crawford,J.T. and Shinnick,T.M.  
AUTHORS The rpoB gene of Mycobacterium tuberculosis  
TITLE Antimicrob. Agents Chemother. 38 (4), 805-811 (1994)  
JOURNAL 94304130  
MEDLINE  
PUBMED 8031050  
FEATURES Location/Qualifiers  
source 1. 5084  
/organism="Mycobacterium tuberculosis"  
/strain="Rv"

/db\_xref="taxon:1773"  
1065. 4598  
/gene="rpoB"  
1065. 4598  
/gene="rpoB"  
/codon\_start=1  
/evidence=experimental  
/transl\_table=1  
/product="RNA polymerase beta-subunit"  
/protein\_id="AA21416.1"  
/db\_xref="GI:468334"  
/translation="MLEGCILADSRQSKTAASPSRPSQSSNNNSVPGAPNRVSFAKL  
REPLEVGLLDVDTGSFELIGSPWRRESAAERGVNPGVGLLEVLVLSPIEDFSGS  
MSLSFSDPRFDVVKAPVDECKDKDMYTAAPLFTVAETNNNTGKTSQVPMGDFPMH  
TEKGTFLINGTERVVSQVSPGVVFDETIDKSDTKLHSVKVIPSRGAWLEFDVVK  
RDTGVGRIDRKRQPVTVLLKALGHTSQIIVERGFSFSEIMRSTLEKONTVGTDEALLD  
IYKLRPGEPPTKESAOQLLENLFKERRDLARVGRYKNNKLGHLHVGEPITTSSTLT  
EDDVATIEYLVRHLEGQTTMTVPGVEVPTETDDIDHFNRRLLRTVGEIQNGIRVG  
MSRMERVRERTTQDVEAITPQTLINIRPVYAAIKETFEFTSOLFQMDQNNPLSGIT  
HKRRSALGPGLSRERAGLEVDRVHPHYGRMCPKETPEGNIGLIGSLSVYARVNP  
FGFIETPYKVVGVSDGIVVLTADSDHVAQANSPIDADGRFVEPRVLRKAG  
EVEYVPSSEVDYMDSPROMYSVATAMIPLEHDDANRALGMANKQARQVPLVSRSEAP  
LVGTGMLRAAIDAATSSQESGVIEEVSADYIIVMDHNGTRRTYRMRKFARNHGTCP  
ANQCPVLDAGDRVEAGQVIADGCTDDGEMALGNLLVAIMPWEGHNYEDAIIILSNRL  
VEDVLTSHIEHEIDARDTKLGAETTRDIPNISDEVLDLDERGIVRIGAEVRDG  
DILGKVTGPKETELTPEERLLRAIFGKAREVRDTSIKVPHGESGKVGIRVFSRED  
EDELPAQVNELVYVVAQKRKISDGLAGRHNGKVGIGKILPVEDMPFLADGTPVDI  
ILNTHGVPRRNIGQILETHLGCWAHSGWKVDAAGVPMNAARLPDELLLEAQAIVS  
TPVFGCAEAELOGLSCTLPNRDGDVLVDAGKAMLPDGRSGEPPFPYVTVGWTIM  
KLHLVDDKIHARSTGPGYSMITQPLGKAGQGGQRFEGMECMAQVAGAYTLQELL  
TIKSDDTVGRVYKYEIVKGENIPEGPPIESFKVLLKELQSLCLNVLSSDGAATL  
REGDEDLERAAANLGINLSRNESASFEDLA"

4641. 5084  
/gene="rpoC"  
4641. 5084  
/gene="rpoC"  
/codon\_start=1  
/transl\_table=1  
/product="RNA polymerase beta'-subunit"  
/protein\_id="AAA21417.1"  
/db\_xref="GI:537608"  
/translation="MLDYNFDELRLGATAEDIRMSYGEVKKPETINYRTLKPKED  
GLFCEKIFGPTDRMECYCGYKVRFGKIGICGVEYTRAKVRMRGMHIELAAPVT  
HIWFKGVPSRLGYLLDLAPKDLKIIYFRAIVTSVDEEMHNEL"

BASE COUNT 969 a 1534 c 1691 g 890 t

ORIGIN

Query Match 100.0%; Score 19; DB 1; Length 5084;  
Best Local Similarity 100.0%; Pred. No. 36;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCAAGGAGAGCGCTACGA 19  
|||||

Db 1966 TCAAGGAGAGCGCTACGA 1984

RESULT 14  
AE006964  
LOCUS  
DEFINITION Mycobacterium tuberculosis CDC1551, section 50 of 280 of the  
complete genome.  
ACCESSION AE006964 AE000516  
VERSION AE006964.1 GI:13880217  
KEYWORDS  
SOURCE Mycobacterium tuberculosis CDC1551.  
ORGANISM Mycobacterium tuberculosis CDC1551  
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
Actinomycetales; Corynebacterineae; Mycobacteriaceae;  
Mycobacterium; Mycobacterium tuberculosis complex.  
1 (bases 1 to 19352)  
REFERENCE Fleischmann,R.D., Alland,D., Eisen,J.A., Carpenter,L., White,O.,  
AUTHORS Peterson,J., DeBoy,R., Dodson,R., Gwinn,M., Haft,D., Hickey,E.,  
Kolony,J.F., Nelson,W.C., Umayam,L.A., Ermolaeva,M.,

Salzberg, S.L., Delcher, A., Utterback, T., Weidman, J., Khouri, H., Gili, J., Mikula, A. and Bishai, W.  
Whole genome comparison of Mycobacterium tuberculosis clinical and laboratory strains

JOURNAL REFERENCE

Unpublished  
2 (bases 1 to 19352)  
Fleischmann, R.D., Alland, D., Eisen, J.A., Carpenter, L., White, O., Peterson, J., DeBoy, R., Dodson, R., Gwinn, M., Haft, D., Hickey, E., Kolonay, J.F., Nelson, W.C., Umayam, L.A., Ermolaeva, M., Salzberg, S.L., Delcher, A., Utterback, T., Weidman, J., Khouri, H., Gili, J., Mikula, A. and Bishai, W.  
Direct Submission

JOURNAL REFERENCE

Submitted (25-Apr-2001) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA

FEATURES

Location/Qualifiers  
1..19352  
/organism="Mycobacterium tuberculosis CDC1551"  
/strain="CDC1551"  
/db\_xref="taxon:83331"  
/note="clinical strain"  
163..3699  
/gene="MT0695"  
163..3699  
/gene="MT0695"  
/note="similar to GB:L27989 GB:L05910 GB:U12205 SP:P47766  
PID:L49992; identified by sequence similarity; putative"  
/codon\_start=1  
/transl\_table=11  
/product="DNA-directed RNA polymerase, beta subunit"  
/protein\_id="AAK44921.1"  
/db\_xref="GI:13880218"  
/translation="MLEGCIADSRQSKTAASPSRPOSSNNNSVPGAPNRYSFAL  
RELPVGLDVTDSFEWLGSPMRSAERGDVNPVGLLEVLVELSPEDFSGS  
MSLSDFRFDNDKAPVDECKDKMTYAPLEVTAEFINNTGKISQTFVMDGFPMM  
TEKGIFLINTERVVSVLPSPGVYFETDKSTDKFLSHSVKVPISRGAWLEFDVDK  
RUTVGRIDRRKQVTVLLKALGWTSEQIVRFNGFSEIMRSTLEKONTVGTDEALID  
IRKLPGPEPKSAQTLLENLFKFKRYDVLAVGRVYKNNKLGHLVGEPTITSSTLT  
EDVATLEYLRHLEGQTTVPVGVVEVDETIDHFGNRLRTVGEIIONQIRVG  
MSRMERVYRMTTQDVEAITPQTLINRPVVAIKFEGFTSOLSPMDONNPLSGIT  
HKRRLSALPGSRAGLEVRDVPVHSGYRMCPIETPEGNIGLIGLSVYARVP  
FCGFIETPVKVVGVSDIYVLTADIEDRHVQAQSPIDAGRFVPEVPLVRRKAG  
EVEYVPSSEVDYVSPQWVSATAMIFLEHDANRGMNQVAPLVRSRAP  
LVGTCHERLAAIDAGVVVAESGVEVSADYITVMDNGTRRTYBMRFARSNHGT  
CANQCPIDVAGDRVBAQVADGCTDGDGEMALKNLLVAIMPWECHNYEDAILSNR  
LVEEDVLTSHIEEIDARDKLGAEITRIDIPNIDEVLADLDERGIVRIGAEVRD  
GDILKVTVPKGETLTPERLLRAIFGEKAREYDTSKVPHGESCKVIGIESRE  
DEDELPAQVNLVRVYVQKRIISGDKLAGHCKGVICKILPVEDMFLAQGTPTVD  
ITLNTHGQAEALGGLSCTLPNRDGDVLDADGKAMLFDRGSEPEFPYPTVYMYI  
MKLHLVDOKTHARSTGYSMITQOPLGGKAQFGGRFGEKMEQWMAQYGAAYTLQEL  
LTIKSDDTVGRVYVYEAIVKGENIPEGIPEFESFKVLLKELQSLCLNVEVLSSDGAIE  
LREGEDLERAAANLGINLSRESASVEDLA"  
3744..7694  
/gene="MT0696"  
3744..7694  
/gene="MT0696"  
/note="similar to SP:P37871; identified by sequence  
similarity; putative"  
/codon\_start=1  
/transl\_table=11  
/product="DNA-directed RNA polymerase, beta-prime subunit"  
/protein\_id="AAK44922.1"  
/db\_xref="GI:13880219"  
/translation="MLDVNFFDELRIGLATAEDITROMSYGEVKKPETINVTLLPKPKD  
GLFEKIFGPTRWECYCGKRVKRFKGIICERCGVEVTRAKVRRRMRHGLIELAAPVT  
HIWFKGVPKRLGLYLLDAPKDLKIIYFAAYVITSVDEMRHNEISTLEAMAVERRK  
AVEQDRGELPRAQKLEADLAELEAGKADARRKVRGGERMQIKDRDAQRELDOR  
LEDIWSFTKLPQLIVDENLRFVDRYGETFGAMGAESIQKLIENFDIDAES  
LRDVRNGKQKRLRLKLVKAAFOQSGNSPMGVDAVPVPELPRMVLQDGR  
FATSDNLDRVNRNRLKRLDLAGPEIIVNNEKRLMESVDALFDNGRRGRVPT  
GPNRPLKSLDLKQGRFRONLLKRVYDGRSVYVVGPKLQHQCGLPKMALE  
LFKPEVMKRLVDLNAQNTKSAKRMVERORPOVDVLEVIAEHPVLLNRAPTLRLG  
IQATEPMLVEKALQLHPLVCEAFNDFDGDQMAVHPLPSAEQAQARILMLSSNNIL

gene  
CDS

gene  
CDS

SPASGRPLAMPRLDMVTGLYLLTTEVPDTEYQPASGDHPETGYVSSPAAEIAAADR  
GVLSVRKILVRLTOLDRPPVEAEELFGHSGWQPGDANMAETTLGRVWFNELLPLGYIP  
FVNQMHKKVQAAIINDLRAEYPMIVVAGTVDKLDAGFYWATRSYTVSNADVLVPP  
RKEILLDHYEERADKVEKQFQFGALNHDNERALVEIKWKEATDEVGQALREHYPDNDP  
LIITVDSGAIGNFTQTRTLAGMGLVTNPKEFIPRPVKSFRGLTVLEFYFINTHGA  
RKLADTLARTADSGYTLRRLVDVSDVIVREHDDQITQKRGIVVLAERAPDCTILRDP  
YETSAYARTLGTDAVDEAGNVIVRGDQDLGAEIDALLAAGITQVKRVSVLTGATST  
GVCATCYGRSMATGKLVDIGEAGVIVAAQSIGEPGTQITMRTFHGGVGEDITGLPR  
VOELFEARVPRGKAPIADTVGRVLEDEREYKIIIVPDGGEVYVYDKISKRLRLV  
FKHEDGSERVLSGDHVEVGQQLMEGSDAPHEVLRVQSPREVQIHLVREVEQVTRAQG  
VSIHDKHIEVIRQMLRRVTIIDSGSTFELPGSLIDRAEFNAENRRVVAEGEFPAAQR  
PVLMLGITKASLATDSWLSAASFQETTRVLTDAAINCRSDKLNGLKENYIIIGKLIPAGT  
GLNRYNIAVQPTTEARAAAYTIPSYEDQIYSPDFGAATGAAPVLDYDYSYR"  
/gene="MT0697"  
complement(7691..8065)  
/gene="MT0697"  
complement(7691..8065)  
/note="identified by Glimmer2; putative"  
/codon\_start=1  
/transl\_table=11  
/product="hypothetical protein"  
/protein\_id="AAK44923.1"  
/db\_xref="GI:13880220"  
/translation="MFDAAALITNPGHAWASAMERSGLLECAGLDEQPFCEFTADKL  
NPDGSSRRVPRRQADGGIATHVERGGQSGGQAGVYVQPMHGFALMAQDRLIHH  
GQOTQNRQAQAFRVRFVCVCSPT"  
complement(8058..9972)  
/gene="MT0698"  
/note="this region contains an authentic frame shift and  
is not the result of a sequencing artifact; identified by  
Glimmer2; putative; conserved hypothetical protein,  
authentic frameshift"  
10167..10925  
/gene="MT0699"  
10167..10925  
/gene="MT0699"  
/note="identified by match to PFAM protein family HMM  
PF01461"  
/codon\_start=1  
/transl\_table=11  
/product="AP endonuclease, family 2"  
/protein\_id="AAK44924.1"  
/db\_xref="GI:13880221"  
/translation="MLGSHVSPDPLAAAEAGDVYQIFLGNPSWKAKPKPRDDAA  
ALKATLPIYVHAPYLINLASANNRVIPSRKILQETCAAADIGAAAVIVHGHVAD  
DNIDKGFQWRKALDRLETPVPLVLENTAGDHDHARRDITARLMDVIGDTGIGFC  
LDTCHTWAAGALTDVADRIKAITGRIDLVCNDRDEAGSGRRHRANLGSQIDPDL  
LVAAVKAAGAPVICETADQGRKDDIAFLRBTGS"  
10957..11799  
/gene="MT0700"  
10957..11799  
/gene="MT0700"  
/note="similar to GB:U00012 PID:466863; identified by  
sequence similarity; putative"  
/codon\_start=1  
/transl\_table=11  
/product="hydrolase/esterase, putative"  
/protein\_id="AAK44925.1"  
/db\_xref="GI:13880222"  
/translation="MLRRVAILLAALVAFAGSGGTRLAAGFCNGNSVHTLOVDGAGR  
SYRLYKPVGLPSSAPLVMLHGGPSAKOERSYGVHDELADSEKFLVAVPDGVRHAWN  
ANGGCCGCRPAREGVDDIGFVRAVDIANNVSDIPARVITVGHNSGAIWSTYLAACN  
SIFAAGVSGTQDPCQSPRVSVIHHGTADPLRVYHGGPGAGFARIDGPPDPLN  
AFHVRNRCGALDTTTEGPTTSGATCADNRNRVLLTVDDAGHRWFSFATQTLWRFEA  
AHFR"  
11859..13487  
/gene="MT0701"  
11859..13487  
/gene="MT0701"  
/note="similar to SP:P33224 GB:L20915 PID:457172  
PID:457174 PID:537028; identified by sequence similarity;  
putative"  
/codon\_start=1

gene  
CDS

gene  
CDS

```

/transl_table=11
/product="acyl-CoA dehydrogenase, putative"
/protein_id="AAK44926.1"
/db_xref="GI:13880223"
/translation="MSDTHVVTNQVPLENYPASSPVLIIEALIQEGGQWGLDEYNEV
WADISGCOARWELADNRNPILTHDAYGYRDEVEYDPAYHELMRTAITHGMAAP
WADISGCOARWELADNRNPILTHDAYGYRDEVEYDPAYHELMRTAITHGMAAP
DEPLKATPATKAGISMTKVEPGHICPISMTYAVVVPALRYNSELAAYVEPLLTREY
IFVLQAQPGSLCFLPVLVLPDSTRNMFLOKDKLGNHANASSEVEYDCAWLV
GEGRGVPTIIEWNLTRLDALGSLTSMRTGLTRAVHQAHRKAFGAYLIDQPLMRN
VLADLAVEARATIVARMAGATDNVAGNETEALLRRIGLAAAYWCKRSTAAAE
ALECLGNGVSDSGMPRLYREAPLNGIWECSNVSALDTRAMATRPACVEVLPDEL
ARSAGDPRLDGHVERLRLQGLDITGYRARKIAEDICLALQGLSLLVRHGHVAAEA
FLATRLGQMGGAYTMPAGLDLAPILERALVKG"
13498..14436
/gene="MT0702"
CDS
13498..14436
/gene="MT0702"
/notes="similar to GP:3885480; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="enoyl-CoA hydratase/isomerase family protein"
/protein_id="AAK44927.1"
/db_xref="GI:13880224"
/translation="MTHAIRPVDNLKMTYEVGTGRIARITFNRPEKNAIADTPL
ELSAVNERADLPVHVILVSGRGEFCAGFDLSAAGSSSTGGGAGVQGTVLDDGT
QAVNHLNPQMPDMIDYQNMRSFRVGFASLMHAKPTVVKIHGYCVAGSTDLAHQD
VIAAADAKIGYPPTRVGVPAAGLWAHRLGDQRAKRLFTGDCITGAQAQWGLAVEA
PEPADIDTERELVRIALPYNQLIWLKLSALLQGVATSRMVSSTVDFGAARHT
PEGHAFVADHVGFRDVARRRDPFGDYGROASRV"
14439..15161
/gene="MT0703"
CDS
14439..15161
/gene="MT0703"
/notes="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
100.0%; Score 19; DB 1; Length 19352;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCAAGGAGAGCGCTACGA 19
|||||
DB 1064 TCAAGGAGAGCGCTACGA 1082

RESULT 15
LOCUS MTC1376 19770 bp DNA linear BCT 03-AUG-2001
DEFINITION Mycobacterium tuberculosis H37Rv complete genome; segment 37/162.
ACCESSION 295972 AL123456
VERSION 295972.1 GI:3261790
KEYWORDS
SOURCE Mycobacterium tuberculosis H37Rv.
ORGANISM Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Corynebacterineae; Mycobacteriaceae;
Mycobacterium; Mycobacterium tuberculosis complex.
1 (bases 1 to 19770)
Cole,S.T., Brosch,R., Parkhill,J., Garnier,T., Churcher,C.,
Harris,D., Gordon,S.V., Eiglmeier,K., Gas,S., Barry III,C.E.,
Tekala,F., Badcock,K., Basham,D., Brown,D., Chillingworth,T.,
Connor,R., Davies,R., Devlin,K., Feltwell,T., Gentles,S.,
Hamlin,N., Holroyd,S., Hornsby,T., Jagels,K., Krogh,A., McLean,J.,
Moule,S., Murphy,L., Oliver,S., Osborne,J., Quail,M.A.,
Rajandream,M.A., Rogers,J., Rutter,S., Seeger,K., Skelton,S.,
Squires,S., Squares,R., Sulston,J.E., Taylor,K., Whitehead,S., and
Barrell,B.G.
Deciphering the biology of Mycobacterium tuberculosis from the
complete genome sequence

```

Nature 393 (6685), 537-544 (1998)  
98295987  
634230  
2 (bases 1 to 19770)  
Parkhill,J.  
Direct Submission  
Submitted (11-JUN-1998) Submitted on behalf of the Mycobacterium tuberculosis sequencing and mapping teams, Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA Unite de Genetique Moleculaire Bacterienne, Institut Pasteur, 28 rue du Docteur Roux, 75724 Paris Cedex 15, France E-mail: parkhill@sanger.ac.uk  
On Jun 27, 1998 this sequence version replaced gi:2143285.  
Notes:  
Details of M. tuberculosis sequencing at the Sanger Centre are available on the World Wide Web.  
(URL, http://www.sanger.ac.uk/Projects/M.tuberculosis/) CDS have been renumbered from the original cosmid submissions but the Old gene designations are in brackets after the new gene numbers.  
Gene prediction was based on a Hidden Markov Model of TB genes implemented in TParse (Krogh) supplemented with visual inspection of positional base preference in codons, especially where there is an increase in the observed/expected third position G + C.  
CAUTION: In some cases we may not have predicted the correct (atg, gtg, or ttg) which is preceded by an upstream ribosome binding site sequence (optimally 5-13bp before the initiation codon). If this cannot be identified we choose the most upstream initiation codon.  
Location/Qualifiers  
1..19770  
/organism="Mycobacterium tuberculosis H37Rv"  
/strain="H37Rv"  
/db\_xref="taxon:83332"  
/clone\_xref="I376"  
57..61  
/note="agga, possible rbs upstream of Rv0654"  
68..1573  
/gene="Rv0654"  
68..1573  
/gene="Rv0654"  
/note="Rv0654", (MTC1376.22), len: 501. unknown, FASTA score: Q53353 LIGNOSTILBENE-ALPHA, BETA-DIOXYGENASE (485 aa) opt:280 z-score: 330.1 E(): 2.3e-11, (28.5% identity in 523 aa overlap). Also similar to M. tuberculosis protein MTCY21C12.07c (29.5% identity in 522 aa overlap)"  
/codon\_start=1  
/transl\_table=11  
/product="hypothetical protein Rv0654"  
/protein\_id="CAB09380.1"  
/db\_xref="GI:2143307"  
/db\_xref="SPTREMBL:O06785"  
/translation="MTTAAAESEONPYLEGLFAPVSTEVATDLPVTGRIPHELDGRY LRGNPNVAEVDPATYTHWFTGDAMVHGVALRDGKARWYRNVRTPAVCAALGEPISA RPHRTGIEGGPNVTLTHAGRTLALVAGVWVVELTDELTVGCDPDLTGHGVT AHPRDPTGHTGELHAVSYSFARGHRYQVSVICTGDGHARTVDIEVAGSPMHNSFLDTN RYVYDLPVTFDPMOVPAVSPRWLQRPALVIOQLVGRVIRPDIPIALNRNQGSHD RLYANWPSYPARVGVMPREGNEDVRFWDIEPCYVYTHPLNAYSECRNGAEVLVLOW RYSRNFDRDRGPGGDSRLDRTINLATGAVTAECRDDRADEFPFRINETLVGGPHR FAYTVIEGGFLVGAGAAALSTPLYKQDCVTGSSVSLDPLDILIGENHVFYVNPSPARAE DDGILMGYGHWRGDEGDEGLLLDQAQTLESATVHLQRPVPMGFHGNWAPT"  
1585..2664  
/gene="Rv0655"  
1585..2664  
/gene="Rv0655"  
/note="Rv0655", (MTC1376.21), len: 359, abc transporter, FASTA score: YRBF\_ECOLI P45393 hypothetical abc transporter atp-binding (269 aa) opt: 644 z-score: 721.8 E(): 3.4e-33 (38.5% identity in 244 aa overlap); contains PS00017 ATP/GTP-binding site motif A, PS00211 ABC transporters family signature, highly similar to M. leprae MKL\_MYCLE P30769 possible ribonucleotide transport atp- (347 aa) opt: 2021 z-score:2244.4 E(): 0, (92.2% identity in 335 aa overlap). Also similar to many otherM.

```

tuberculosis ABC transporters eg. MTCY253.24 (33.6%
identity in 241 aa overlap)"
/codon_start=1
/transl_table=11
/product="hypothetical protein Rv0655"
/protein_id="CAB09379.1"
/db_xref="GI:2143306"
/db_xref="SWISS-PROT:O06784"
/translation="MRYSDSYHTGRQPRASTEGFPMGVSYEVNGLTAKSFGSSRIWE
DYLTLPAGEVSVLGPGSGTKSVFLKSLGLRPGRGSIIDGTDIIKESAKELYEI
RLFLVLFQDGLGSMNLNDYNTAPFLREHTKKKESEIRDIVMEKALVGLGDKERKF
PGEISGMRKAGLARALVLDPQIILCDEPDSGLDPVTAYLSQLIMDINAQIDATIL
IVTHNINARTVPONMGLFRKHLVMEFGPEVRLTSDPVVROFLNGRRIGPIGMSEE
KDEATMAEEQALLDAGHHAGGVEIEGVPPQISATPCMPERKAVARQARVREMLHTL
PKKAQAAILDDLEGTHKYAVHICQ"
misc_feature
1762..1785
/gene="Rv0655"
/note="PS00017 ATP/GTP-binding site motif A"
2074..2118
/gene="Rv0655"
/note="PS00211 ABC transporters family signature"
complement(3052..3435)
/gene="Rv0656c"
complement(3052..3435)
/gene="Rv0656c"
/note="Rv0656c" (MTCI376.20, unknown), len: 127 aa"
/codon_start=1
/transl_table=11
/product="hypothetical protein Rv0656c"
/protein_id="CAB09378.1"
/db_xref="GI:2143305"
/db_xref="SPTREMBL:O06783"
/translation="MAAATTGTHRGLELRAAORAVGSCPEPQAEFCRSARNADEFDQ
MSRMFGDVPDPVPKSVMRWIDSQOHLRAGAVGALSVDLLICDTAAAGLVLVHL
DDADYELAEHLDIRRVVSADD"
complement(3530..3685)
/gene="Rv0657c"
complement(3530..3685)
/gene="Rv0657c"
/note="Rv0657c" (MTCI376.19), unknown, len: 51 aa; similar
to several other M. tuberculosis hypothetical proteins eg.
YW08_MYCTU Q10848 hypothetical 8.9 kd protein cy39.08c (80
aa), fasta scores; opt: 107 z-score: 182.3 E(): 0.0038,
45.8% identity in 48 aa overlap. Also similar to MTCY48.5
andAL020958|SC4H8.7 Streptomyces coelicolor cosmid 4H8 (66
aa). 41.0% identity in 39 aa overlap."
/codon_start=1
/transl_table=11
/product="hypothetical protein Rv0657c"
/protein_id="CAB09377.1"
/db_xref="GI:2143304"
/db_xref="SPTREMBL:O06782"
/translation="MSVTQIDLDLADVMRTAAVHTKKEAVNLAMRDYVERFRRIE
ALARSRE"
complement(3691..3696)
/note="possible RBS upstream of Rv0657c"
complement(3761..4477)
/gene="Rv0658c"
complement(3761..4477)
/gene="Rv0658c"
/note="Rv0658c" (MTCI376.18), len: 238, unknown,
probablemembrane protein, similar to YPRB_ECOLI P33774
hypothetical24.3 kd protein (urf 1) (217 aa), fasta
scores; opt: 174 z-score: 215.8 E(): 5.3e-05, (25.6%
identity in 223 aa overlap). Also similar to MTCY359.10
(28.7% identity in 178 aa overlap)"
/codon_start=1
/transl_table=11
/product="hypothetical protein Rv0658c"
/protein_id="CAB09376.1"
/db_xref="GI:2143303"
/db_xref="SPTREMBL:O06781"
/translation="MEAGRADTVAPSHRWGLGAFVLVFLVASTSLAVVLTGHGPV
SAGVLALALAAPTVVAAGLAILITRLRNGNLTDLRLRWSRGLRLGLMFGFGMLVT
IPASLVYTAIVGPEANSVAVRIFGCVASPMFALVFLVVFVAPLCEEIYRGLLWG
AVDI:RWGRWAALVVTVVFAHLAHLFAFARAPLLVVAIPIALARYSGLLASIVTHQV
TNLLPGIIVLLGLTGATSLP"
complement(4480..4483)
/note="possible RBS upstream of Rv0658c"
complement(4753..5061)
/gene="Rv0659c"
complement(4753..5061)
/gene="Rv0659c"
/note="Rv0659c" (MTCI376.17), len: 102; unknown, similar
to YW28_MYCTU Q10867 hypothetical 12.3 kd protein cy39.28
(114 aa), fasta scores; opt: 144 z-score: 213.2 E():
7.3e-05, 30.8% identity in 107 aa overlap. Also similar to
MTCY09F9.22 (32.7% identity in 101 aa overlap)"
/codon_start=1
/transl_table=11
/product="hypothetical protein Rv0659c"
/protein_id="CAB09387.1"
/db_xref="GI:2143302"
/db_xref="SPTREMBL:O06780"
/translation="MRRGELWFAATPGGDRPVLVLTDRDPVADRIGAVVVVALTRTRRG
LVSELELTAVENRVPSDCVYVNFQNIHTLPRTAFRRITRLSPARLHEACQTLRASTGC
"
complement(5048..5293)
/gene="Rv0660c"
complement(5048..5293)
/gene="Rv0660c"
/note="Rv0660c" (MTCI376.16), len: 81, some similarity to
IAF016485.130 Halobacterium sp: NRC-1 plasm (100 aa),
32.4% identity in 74 aa overlap"
/codon_start=1
Query Match 100.0%; Score 19; DB 1; Length 19770;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TCAAGGAGAGCGCTACGA 19
|||||
DB 10758 TCAGGAGAGCGCTACGA 10776
Search completed: November 13, 2002, 01:28:14
Job time : 73.7983 secs

```

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 12, 2002, 16:59:42 ; Search time 1109.13 Seconds  
(Without alignments)  
3037.202 Million cell updates/sec

Title: US-09-697-123B-22

Perfect score: 208  
Sequence: 1 lcaaggaagacgcctacgac.....ccggtgaacacgcagacat 208

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 1615406 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST: \*  
1: em\_estba: \*  
2: em\_esthm: \*  
3: em\_estlin: \*  
4: em\_estlm: \*  
5: em\_estmv: \*  
6: em\_estol: \*  
7: em\_estro: \*  
8: em\_hic: \*  
9: gb\_est1: \*  
10: gb\_est2: \*  
11: gb\_hic: \*  
12: gb\_est3: \*  
13: gb\_est4: \*  
14: gb\_est5: \*  
15: em\_estfun: \*  
16: em\_estcom: \*  
17: gb\_gss: \*  
18: em\_gss\_hum: \*  
19: em\_gss\_inv: \*  
20: em\_gss\_pln: \*  
21: em\_gss\_vrt: \*  
22: em\_gss\_fun: \*  
23: em\_gss\_mam: \*  
24: em\_gss\_mus: \*  
25: em\_gss\_pro: \*  
26: em\_gss\_rod: \*  
27: em\_gss\_rod: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	150.4	72.3	1282	9 A1770311	A1770311 42 Mycob
C 2	43.4	20.9	877	17 A0687771	A0687771 nxb0075L
C 3	42.2	20.3	568	13 B1995346	B1995346 1031026E0
C 4	39	18.8	326	10 BE359066	BE359066 Dg1_38_E0
C 5	39	18.8	444	12 BE918452	BE918452 OV1_4_A04
C 6	39	18.8	520	13 B1246073	B1246073 IPL_68_D0

7	39	18.8	528	13 B1074353	B1074353 IPL_14_E0
8	39	18.8	529	10 BE358187	BE358187 Dg1_26_C0
9	39	18.8	535	12 BE918899	BE918899 FMI_1_F11
10	39	18.8	537	12 BG412695	BG412695 OV2_33_F0
11	39	18.8	537	12 BG649762	BG649762 RH12_83
12	39	18.8	550	12 BE917902	BE917902 OV1_7_E07
13	39	18.8	552	10 AW922583	AW922583 Dg1_45_E1
14	39	18.8	555	12 BE918453	BE918453 OV1_4_A05
15	39	18.8	557	12 BG605810	BG605810 RH12_81
16	39	18.8	558	12 BE918353	BE918353 OV1_2_E10
17	39	18.8	562	12 BE918202	BE918202 OV1_3_E09
18	39	18.8	564	12 BE918401	BE918401 OV1_3_B09
19	39	18.8	570	12 BG239936	BG239936 OV1_30_H1
20	39	18.8	571	12 BE918234	BE918234 OV1_3_G04
21	39	18.8	572	12 BE918608	BE918608 OV1_4_A05
22	39	18.8	579	12 BE918149	BE918149 OV1_2_E10
23	39	18.8	580	12 BG049324	BG049324 OV1_18_E0
24	39	18.8	583	12 BE918122	BE918122 OV1_2_C01
25	39	18.8	597	12 BE911885	BE911885 OV2_39_D1
26	39	18.8	599	12 BE917926	BE917926 OV1_16_B0
27	39	18.8	619	13 BM317840	BM317840 OV1_16_B0
28	39	18.8	620	12 BE918695	BE918695 FMI_1_D08
29	39	18.8	623	12 BG049396	BG049396 OV1_19_D0
30	39	18.8	625	12 BG048093	BG048093 OV1_27_E1
31	39	18.8	628	12 BE918704	BE918704 FMI_1_E11
32	39	18.8	632	12 BE918690	BE918690 FMI_1_D03
33	39	18.8	633	12 BG605900	BG605900 RH12_82
34	39	18.8	636	12 BG048276	BG048276 OV1_16_B0
35	39	18.8	636	12 BG356764	BG356764 OV2_9_F09
36	39	18.8	637	12 BG412523	BG412523 OV2_35_D0
37	39	18.8	637	12 BG412409	BG412409 OV2_33_F0
38	39	18.8	646	10 BE356124	BE356124 Dg1_122_H
39	39	18.8	647	12 BE917821	BE917821 OV1_7_E07
40	39	18.8	665	12 BG048108	BG048108 OV1_27_G0
41	39	18.8	700	12 BG605928	BG605928 RH12_83
42	38.6	18.6	489	12 BG047785	BG047785 OV1_27_G0
43	38.4	18.5	772	12 BE618004	BE618004 HVSMEC000
44	38.2	18.4	502	13 B1074667	B1074667 IPL_14_E0
45	37.4	18.0	499	10 AW922389	AW922389 Dg1_18_H0

# ALIGNMENTS

RESULT 1  
A1770311/c  
LOCUS 1282 bp mRNA linear EST 24-JAN-2000  
DEFINITION 42 Mycobacterium anaerobic stationary phase library Mycobacterium  
smegmatis cDNA, mRNA sequence.  
ACCESSION A1770311  
VERSION A1770311.1 GI:6742680  
KEYWORDS EST.  
SOURCE Mycobacterium smegmatis.  
ORGANISM Mycobacterium smegmatis  
Bacteria; Actinobacteria; Mycobacteriaceae; Mycobacterium.  
Corynebacterineae; Mycobacteriales.  
REFERENCE  
1 (bases 1 to 1282)  
Murgas-Oel, B., Tay, A. and Dick, T.  
Upregulation of stress response genes and ABC transporters in  
anaerobic stationary-phase Mycobacterium smegmatis  
Mol. Gen. Genet. 262 (4-5), 677-682 (1999)  
20092472  
JOURNAL MEDLINE  
COMMENT Contact: Murgas-Oel, B.  
Mycobacterium Laboratory  
Institute of Molecular and Cell Biology  
30 Medical Drive, Singapore 117609, Republic of Singapore  
Tel: 65 874 3011  
Fax: 65 779 1117  
Email: mcbomelmb nus.edu.sg  
Insert Length: 1282 Std Error: 0.00  
Seq primer: T3 Forward; T7 Backward.  
Location/Qualifiers  
1..1282

FEATURES  
source

/organism="Mycobacterium smegmatis"  
 /strain="mc215"  
 /db xref="taxon:1772"  
 /clone\_lib="Mycobacterium anaerobic stationary phase  
 library"  
 /lab\_host="E. coli XL1-Blue MRF"  
 /note="Vector: Lambda Zap II; Bacilli were disrupted using a  
 Mini bead beater (Biospec). RNA was isolated using the  
 RNeasy Protocol (Qiagen). Purified RNA was subjected to 2  
 rounds of digestion with RNase-free Dnase I (Promega).  
 Dnase I was heat-inactivated at 75°C for 5 min. and  
 removed by using RNeasy columns followed by phenol  
 extraction and ethanol precipitation. The RNA  
 preparations were confirmed to be free of genomic DNA  
 contamination by carrying out PCR and RT-PCR using the  
 Access kit (Promega) and primers specific for the  
 histone-like protein gene hlp (Lee et al., 1998). cDNA  
 was synthesized using random hexamer primers (Promega)  
 and Stratagene's cDNA synthesis kit. cDNA fragments were  
 ligated into lambda Zap II vector and packaged in vitro  
 using Stratagene reagents."

Query Match	72.3%	Score 150.4	DB 9	Length 1282
Best Local Similarity	82.7%	Pred. No. 2e-29		
Matches 172	Conservative	0	Mismatches 36	Indels 0
			Gaps	0
0y	1 TCAAGCAGAACGGCTTACGACCTGGCCCGCTGGCTGCTATTAAGGTTCAACAGAAAGCTTC	60		
Db	236 TCAAGGGAAGACGGCTTCGACCTGGCCCGCTGGCCGCTTCAAGGTTCACAGAAAGACTGG	177		
0y	61 GCGTCGATGTCGGCGCGCCCATACGTCGTGTGAGCGTCGACCGAAGAACGTCGTGGCCA	120		
Db	176 GCCTGAACCGCGGCAACGCCGATACACCAAGCTGACGCTGACCGAAGAGAGCTGTCCGA	117		
0y	121 CCATTCGAATATCTGGTCGCGCTTGCACGAGGGTTCAGACACGATGATCTTCCGGCGCG	180		
Db	116 CCATTCGAGTACTGGTGGCGCTGCGACGAGGGGTCAAGTTCGATGAGACCGTCCGGGTGGCG	57		
0y	181 TTCAGGTCGCGGTCGGAACCGACGACAT	208		
Db	56 TTCAGGTCCCGTCAAGGTGTCAGACAT	29		

RESULT 2	LOCUS	DEFINITION	ACCESSION
AQ687771/c	AQ687771	877 bp	DNA
	nbx00075L01f	CUGI Rice BAC library	Oryza sativa genomic clone
	nbx00075L01f	DNA sequence.	
	nbx007771		

ACCESSION	AK068777.1	GI:5328939
VERSION	AK068777.1	
KEYWORDS	GSS.	
SOURCE	Oryza sativa.	
ORGANISM	Oryza sativa	
	Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta: Spermatophyta: Magnoliophyta, liliopsida: Poales: Poaceae; Ehrhartoideae: Oryzae; Oryza.	
REFERENCE	1 (bases 1 to 877)	
AUTHORS	Wing, R.A. and Dean, R.A.	
TITLE	A BAC End Sequencing Framework to Sequence the Rice Genome	
JOURNAL	Unpublished (1998)	
COMMENT	Contact: wing RA	

FEATURES

Email:	rwj@eng.lemson.edu
Seq primer:	TAATATGACCTCATTATACCG
Class:	BAC ends
High quality sequence stop:	85.
Location/Qualifiers	

```

1. 877
/organism="Oryza sativa"
/strain="Japonica"
/cultivar="Nipponbare"
/db_xref="taxon:4530"
/clone="_nxb0075101f"
/clone_lib="CUCI Rice BAC Library"
/tissue_type="Leaf"
/lab_host="E. coli DH10B"
/notes="Vector: paeloBAC11; Site 1: HindIII; Site 2:
HindIII; Rice is one of the two most popular grains in the
world. Half of the world population especially those
inhabiting highly populated areas of the humid tropics
and subtropics, rely on rice as their primary source of
carbohydrate. Monocotyledonous rice is a diploid plant
(2n=24) with a haploid genome equivalent of 431 Mbp
(Arunaganathan and Earle, 1991). The relatively small
genome of rice, three times larger than that of
Arabidopsis, makes it suitable for genomic studies. In
order to facilitate positional cloning, physical mapping
and genome sequencing of rice, we have constructed a BAC
library from Oryza sativa, Nipponbare variety. The
library contains 36,864 clones with an average insert size
of 128.5 kb providing 10.9 haploid genome equivalents. The
deep coverage allows the isolation a particular sequence
with a probability of 99.9 %. Two high density filters,
each containing 18,432 clones (doubly spotted), represent
the whole library for colony screening."

```

[illegible][illegible]

DCMB Box 91000  
Duke University  
Durham, NC 27708-1000  
Tel: 919 613 8159  
Fax: 919 613 8177

FEATURES  
source

Email: chauser@duke.edu.  
Location/Qualifiers

1..568  
/organism="Chlamydomonas reinhardtii"  
/strain="CC-1690 wild type mt+ 21gr"  
/db\_xref="taxon:3055"  
/clone\_lib="C. reinhardtii CC-1690, Stress II (normalized  
) lambda zap II"

/note="Vector: Bluescript II SK-; Site 1: EcoRI; Site 2: XhoI; Stress condition II library, constructed by John Davies and Jeffrey McDermott, combines cDNAs from CC-1690 cells grown to mid-log phase in TAP (NH4+ - containing) and shifted to TAP - NO3- (24hrs); H2 production conditions (0, 12hr, 24hr) see Mellis et al., (2000) Plant Phys. 122: 127-135; TAP + H2O2 (1, 12, 24 hr); TAP + sorbitol (1, 2, 6, 24 hr); TAP + Cd (1, 2, 6, 24 hr). PolyA mRNA was purified from each sample, pooled and cDNA synthesized. The cDNA was directionally cloned into lambda zap II (Stratagene) in the EcoRI (5') and XhoI (3') sites. pBluescript II SK- plasmids were excised from the lambda zap clones by superinfection with Exassist (Stratagene) phage. The library was normalized using method 4 described in Bonaldo et al., (1996) Genome Research 6: 791-806."

BASE COUNT 80 a 206 c 185 g 97 t

ORIGIN

Query Match 20.3%; Score 42.2; DB 13; Length 568;  
Best Local Similarity 50.2%; Pred. No. 0.4;  
Matches 104; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

OY 2 CAAGGAGAGCGCTGACCTGCGCCGCTGCTGCTATAGGTCAACAAGACTCG 61  
DB 251 CAAGGAGAGCGCTGACCTGCGCCGCTGCTGCTATAGGTCAACAAGACTCG 310  
OY 62 GCTGATGTCGGCGAGCCCATCATCATGCTGACGCTGACGGAAGAGCTGCGCCAC 121  
DB 311 CGTCCGCTACAGGAGGAGCCCGCGGCTCTCCGCGCCGCTGCTACCACTGCGCGC 370  
OY 122 CATCGAATATCTGTCCTGCTGACAGAGGTCAAGACCATGATGCTTCGGGCGCGCT 181  
DB 371 CGCGCGCGAGCCCGCGCTGCTCCCGCGGCGCGAGCGCGGTTCGCTCCACCTCGCG 430  
OY 182 CGAGTGTCCGCTGCAACCGACGACAT 208  
DB 431 CGACAGCGCCGCTGCGCTGCTGCGCAT 457

RESULT 4  
BE359066

LOCUS DGI\_38.E06.g1\_A002 Dark grown 1 (DGI) Sorghum bicolor cDNA, mRNA  
DEFINITION sequence.

ACCESSION BE359066  
VERSION BE359066.1 GI:9300623  
KEYWORDS EST.

SOURCE

ORGANISM

Sorghum bicolor  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC  
clade; Panicoideae; Andropogoneae; Sorghum.  
1 (bases 1 to 326)  
Cordonnier-Pratt,M.-M., Gingle,A., Marsala,C., Sudman,M. and Pratt  
,L.H.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

An EST database from Sorghum: dark-grown seedlings  
Unpublished (2000)  
Contact: Cordonnier-Pratt MM  
Laboratory for Genomics and Bioinformatics  
The University of Georgia, Department of Plant Biology  
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
Tel: 706 542 1860  
Fax: 706 583 0210  
Email: mmp@pratt@uga.edu

Sequences have been trimmed to exclude PolyA, vector and regions below Phred quality 16. The threshold for highest quality sequence is 20.  
Seq primer: PolYTmX  
High quality sequence start: 5  
High quality sequence stop: 324  
POLYA=yes.

FEATURES  
source

Location/Qualifiers  
1..326  
/organism="Sorghum bicolor"  
/db\_xref="taxon:4558"  
/clone\_lib="Dark Grown 1 (DGI)"

/note="Organ: 5-day-old dark-grown seedlings; Vector: lambda zap; Site 1: XhoI; Site 2: EcoRI; The library was made from poly-A RNA in the cloning vector lambda zap II. Clones to be sequenced were prepared by mass excision."  
BASE COUNT 72 a 93 c 118 g 43 t

ORIGIN

Query Match 18.8%; Score 39; DB 10; Length 326;  
Best Local Similarity 50.3%; Pred. No. 2.5;  
Matches 96; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

OY 2 CAAGGAGAGCGCTACGACCTGCGCCGCTGCTGCTATAGGTCAACAAGACTCG 61  
DB 117 CAAGTGTGTCGTACAGCTGCTGAGAGCGGAGCTGCTTACTTACAAAGACTTCAA 176  
OY 62 GCTGATGTCGGCGAGCCCATCATCATGCTGACGCTGACGCAAGAGCTGCGCCAC 121  
DB 177 GATCAGCTGAAGGTGCGAGGCGCGCGCGGCGGAGCGGCGGTAGCGCGCGCGCT 236  
OY 122 CATCGAATATCTGTCCTGCTGACAGAGGTCAAGACCATGATGCTTCGGGCGCGCT 181  
DB 237 GCTCAATGTCATATGAGCTTCAGCAAGCCAGACAGCTGCTCCAGCCGAGCTCAT 296  
OY 182 CGAGTGTCCG 192  
DB 297 CAAGGAGACG 307

RESULT 5  
BE918452

LOCUS OV1\_4.A04.g1\_A002 Ovary 1 (OV1) Sorghum bicolor cDNA, mRNA  
DEFINITION sequence.

ACCESSION BE918452  
VERSION BE918452.1 GI:10421570  
KEYWORDS EST.

SOURCE

ORGANISM

Sorghum bicolor  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC  
clade; Panicoideae; Andropogoneae; Sorghum.  
1 (bases 1 to 444)  
Cordonnier-Pratt,M.-M., Gingle,A., Marsala,C., Sudman,M. and Pratt  
,L.H.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

An EST database from Sorghum: ovaries of varying immature stages  
Unpublished (2000)  
Contact: Cordonnier-Pratt MM  
Laboratory for Genomics and Bioinformatics  
The University of Georgia, Department of Plant Biology  
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
Tel: 706 542 1860  
Fax: 706 583 0210  
Email: mmp@pratt@uga.edu  
Sequences have been trimmed to exclude PolyA, vector and regions below Phred quality 16. The threshold for highest quality sequence is 20.  
Seq primer: PolYTmX  
High quality sequence start: 90  
High quality sequence stop: 427  
POLYA=no.

FEATURES

Location/Qualifiers

```

source
1. .444
/organism="Sorghum bicolor"
/db_xref="Taxon:4558"
/clone_1fb="Ovary 1 (OV1)"
/notes="Organ: Mix of ovaries of varying immature stages
from 8-week-old plants; Vector: pBluescript II from Lambda
zap II; Site_1: XhoI; Site_2: EcoRI; The library was made
from poly-A RNA in the cloning vector lambda Zap II.
Clones to be sequenced were prepared by mass excision."
BASE COUNT      102 a      131 c      142 g      69 t
ORIGIN

```

	Query Match	18.8%	Score 39:	DB 12:	Length 444:
	Best Local Similarity	50.3%	Pred. No.	2.6:	
	Matches	96:	Conservative	0:	Mismatches 95; Indels 0; Gaps 0;
OY	2 CAAAGCAACGCCCTACGACCCTGGCCCGGTGCCTATTAAGTCAACAAGAAGCTCGG	61			
Db	162 CAAGGTGTGTTCTACTTGCGTGTGAAGGAGAGCTGATTCTTCAACAAGACTTCAA	221			
OY	62 GCTCATGTGGGGAGCACCCATCAGTCGTCGACGGCTGCACGAAGAAGACTGCTGGCCAC	121			
Db	222 GATACCGCTAAGAGTGGCCCCAAGCGCCGGCCGGCGGAGGAGGCGAAGGCTTAGGCCCCCGCT	281			
OY	122 CATCGAATAATCTGTGTCGCTTGSCACAGAGGTCAAGCCACGATGATCTTGGCGGCGCGT	181			
Db	282 CGCTAACGTGGCCATGAGATTGCAGCAAAGGCCAAGACGAGGTGCCCGACCGGACGTCAT	341			
OY	182 CGAAGTCCCG 192				
Db	342 CAAAGACCCG 352				

RESULT	6
B1246073	
LOCUS	
DEFINITION	B1246073                  520 bp       mRNA       EST 17-JUL-2001 IPL_68.D09.bL.A002 Immature pannicle 1 (IPI) Sorghum bicolor cDNA, mRNA sequence.
ACCESSION	B1246073
VERSION	B1246073.1
KEYWORDS	GI:14824091
SOURCE	EST.
ORGANISM	sorghum. Sorghum bicolor Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophytas; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC Clade; Panicoideae; Andropogoneae; Sorghum. 1 (bases 1 to 520) L.H., R.R., Cordonnier.Pratt.M., M., Gingle,A., Sudman,M. and Pratt .L.H.
REFERENCE	An EST database from Sorghum: developing preanthesis panicles Unpublished (2001)
AUTHORS	Contact: Cordonnier-Pratt MM Laboratory for Genomics and Bioinformatics The University of Georgia, Department of Plant Biology Plant Sciences Building, Km. 2502, Athens, GA 30602-7271, USA Tel: 706 542 1860 Fax: 706 583 0210 Email: mmp@prattuga.edu
TITLE	Sequences have been trimmed to exclude POLYA, vector and regions below Phred quality 16. The threshold for high quality sequence is 20. Three-prime sequences, which are obtained with PolyTMix or T7 Seq primer: JEN REV High quality sequence stop: 400 POLYA-No.
JOURNAL COMMENT	

```

FEATURES
source
    location/Qualifiers
    1. 520
    /organism="Sorghum bicolor"
    /cultivar="Brx623"
    /db_xref="taxon:4558"
    /clone_11b="Immature panicle 1 (Ip1)"
    /note="Organ: Developing preanthesis panicles; Vector
    bluescript II SK(-) from Lambda Zap II; Site:1: XhoI;

```

```

Query Match      18.8%      Score 39;  DB 13;  Length 520;
Best Local Similarity 50.3%      Pred. No. 2.7;
Matches 96;  Conservative 0;  Mismatches 95;  Indels 0;  Gaps 0;

BASE COUNT      113 a      153 c      178 g      76 t
ORIGIN

```

QY	2	CAAGAGAAACGGCTACGACCTGGACCCGGGTGGGTCGCTATAAGGTCAACAAAGAACTCGG	61
Db	300	CAAGGTGGTGTCTACAAGCGTGGTGAACGGCAGCTGGTTGACTTCTACAAAGAACTTCAA	359
QY	62	GCTGCATGTCGGCGGAGCCCATCAGCTCGTGCAGCCTGACCGAAGAAAGAGCTGTGGCCAC	121
Db	360	GATCACGCTGAAGGTGGCCAAAGCGCGCGCGGAGGGAAGGTGAGGCCGCGCCCT	419
QY	122	CATCGAATATCTGTGCTCCGCTTGACACGAGGGTCAACACACATCATCTGTTCCGGCGGCGT	181
Db	420	CGTAACTGGTTCATGAGATTGCACAAAGCCAAACGACCAAGTGTCCCGACCCGAGCTAT	479
QY	182	CGAGGTGCCG	192
Db	480	CAAGAGAACCG	490

RESULT 7	
LOCUS	BI074353
DEFINITION	BI074353 528 bp mRNA linear EST 20-JUN-2001
ACCESSION	IP1_14_E03.g1.A002 Immature panicle 1 (IP1) Sorghum bicolor cDNA,
VERSION	BI074353
KEYWORDS	EST.
SOURCE	BI074353.1 GI:14513010
ORGANISM	Sorghum bicolor

REFERENCE AUTHORS	TITLE JOURNAL COMMENT
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoidae; Andropogoneae; Sorghum. 1 (bases 1 to 528)	
Klein, R.R., Cordomier-Pratt, M.-M., Gingle, A., Sudman, M. and Pratt, L.H.	An EST database from Sorghum: developing preanthesis panicles Unpublished (2001) Contact: Cordomier-Pratt MM

Laboratory for Genomics and Bioinformatics  
The University of Georgia, Department of Plant Biology  
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
Tel.: 706 542 1860  
Fax: 706 583 0210  
Email: [mpmrpratt@uga.edu](mailto:mpmrpratt@uga.edu)  
Sequences have been trimmed to exclude PolyA, vector and regions below Phred quality 16. The threshold for high quality sequence is 20. Three-prime sequences, which are obtained with PolyMx or T7 sequencing primer, are presented as the reverse complement.  
Seq primer: 77  
High quality sequence start: 13  
High quality sequence stop: 528  
PolyA=Yes.

FEATURES	source
Location/Qualifiers	1. .528
/organism="Scorpium bicolor"	/cultivar="BRX623"
/db_xref="taxon:4558"	/clone_id="Immature panicle 1 (P1)"
/note="Organ: Developing preanthesis panicles; Vector: pBluescript II SK(-) from Lambda zap II; Site: 1' xho; Site_2: EcoRI; The library was made from poly-A RNA in the cloning vector lambda zap II. Clones to be sequenced were prepared by mass excision."	
BASE COUNT	103 a 147 c 164 g 114 t



512 CAAGGTCGGTCTACAGCCTGGCGGCAGCCGTGACTCTACAGAAGCTCAA 371

368 GATCACCCTGAAGGTGGCCAAAGCGGCGCGGAGGGCGAGGTGAGGCGCGCGCCGT 427

OY 122 CATGCAATATCTGTCCTCCGCTTGACAGAGGTCAGACACGATGATCGTCCGGCGCGT 181  
 DB 428 CGTCAACTGTCATGAGCTTCGACAAAGCCACAGCAGCAGGTGCCCAACCGGACGTCAT 487  
 OY 182 CGAGTGCCTGG 192  
 DB 488 CAAGGAGACCG 498

RESULT 10  
 BC412695 537 bp mRNA linear EST 13-MAR-2001  
 LOCUS OV2\_33\_f06.b1\_A002 Ovary 2 (OV2) Sorghum bicolor cDNA, mRNA  
 DEFINITION sequence.  
 ACCESSION BC412695  
 VERSION BC412695.1 GI:13318248  
 KEYWORDS EST.  
 SOURCE Sorghum.  
 ORGANISM Sorghum bicolor

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC  
 clade; Panicoideae; Andropogoneae; Sorghum.  
 1 (bases 1 to 537)  
 Cordonnier-Pratt,M.-M., Gingle,A., Marsala,C., Sudman,M. and Pratt  
 L.H.

TITLE An EST database from Sorghum: ovaries of varying immature stages  
 JOURNAL Unpublished (2000)  
 COMMENT Contact: Cordonnier-Pratt MM  
 Laboratory for Genomics and Bioinformatics  
 The University of Georgia, Department of Plant Biology  
 Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
 Tel: 706 542 1860  
 Fax: 706 583 0210  
 Email: mmp@prattuga.edu  
 Seq primer: JEN REV  
 High quality sequence stop: 485  
 POLYA-No.

## FEATURES

source Location/Qualifiers  
 1..537  
 /organism="Sorghum bicolor"  
 /db\_xref="taxon:4558"  
 /clone\_lib="Ovary 2 (OV2)"  
 /note="Organ: Mix of ovaries of varying immature stages  
 from 8-week-old plants. Vector: pBluescript II from Lambda  
 Zap II; Site\_1: XhoI; Site\_2: EcoRI; The library was made  
 from poly-A RNA in the cloning vector lambda Zap II.  
 Clones to be sequenced were prepared by mass excision."  
 BASE COUNT 117 a 159 c 183 g 78 t  
 ORIGIN

Query Match 18.8%; Score 39; DB 12; Length 537;  
 Best Local Similarity 50.3%; Pred. No. 2.7;  
 Matches 96; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

OY 2 CAAGAGAAGCCCTACGACCTGCGCCGCGTGGCTATTAAGGTCAACAAAGAGCTCG 61  
 DB 306 CAAGGTGCTGTCATGACGCTGTGACGCGGAGCTGTTGACTTCAAGAACTTCAA 365  
 OY 62 GCTGATGTCGGCGAGCCCATCATCGTCGACGCTGACCGAAGAAGAGCTGTGGCCAC 121  
 DB 366 GATCAGCGCTGAAGGTGGCCAAAGCGCGCGGAGGCGAGGGGTAGAGCGCGGCCCT 425  
 OY 122 CATGCAATATCTGTCCTCCGCTTGACAGAGGTCAGACACGATGATCGTCCGGCGCGT 181  
 DB 426 CGTCAACTGTCATGAGCTTCGACAAAGCCACAGCAGCAGGTGCCCAACCGGACGTCAT 485  
 OY 182 CGAGTGCCTGG 192  
 DB 488 CAAGGAGACCG 496

RESULT 11

BG649762  
 LOCUS BG649762 537 bp mRNA linear EST 24-APR-2001  
 DEFINITION RH12.83\_c10.b1\_A003 Rhizome2 (RH12) Sorghum prolinguam cDNA, mRNA  
 sequence.  
 ACCESSION BG649762  
 VERSION BG649762.1 GI:13784874  
 KEYWORDS EST.  
 SOURCE Sorghum prolinguam.  
 ORGANISM Sorghum prolinguam

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC  
 clade; Panicoideae; Andropogoneae; Sorghum.  
 1 (bases 1 to 537)  
 Cordonnier-Pratt,M.-M., Gingle,A., Paterson,A., Sudman,M. and Pratt  
 L.H.

TITLE An EST database from Sorghum: Sorghum prolinguam rhizomes  
 JOURNAL Unpublished (2000)  
 COMMENT Contact: Cordonnier-Pratt MM  
 Laboratory for Genomics and Bioinformatics  
 The University of Georgia, Department of Plant Biology  
 Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
 Tel: 706 542 1860  
 Fax: 706 583 0210  
 Email: mmp@prattuga.edu  
 Seq primer: JEN REV  
 High quality sequence stop: 396  
 POLYA-No.

## FEATURES

source Location/Qualifiers  
 1..537  
 /organism="Sorghum prolinguam"  
 /db\_xref="taxon:132711"  
 /clone\_lib="Rhizome2 (RH12)"  
 /note="Organ: Rhizomes; Vector: pBluescript II from Lambda  
 Zap II; Site\_1: XhoI; Site\_2: EcoRI; The library was made  
 from poly-A RNA in the cloning vector lambda Zap II.  
 Clones to be sequenced were prepared by mass excision."  
 BASE COUNT 119 a 158 c 181 g 79 t  
 ORIGIN

Query Match 18.8%; Score 39; DB 12; Length 537;  
 Best Local Similarity 50.3%; Pred. No. 2.7;  
 Matches 96; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

OY 2 CAAGAGAAGCCCTACGACCTGCGCCGCGTGGCTATTAAGGTCAACAAAGAGCTCG 61  
 DB 296 CAAGGTGCTGTCATGACGCTGTGACGCGGAGCTGTTGACTTCAAGAACTTCAA 355  
 OY 62 GCTGATGTCGGCGAGCCCATCATCGTCGACGCTGACCGAAGAAGAGCTGTGGCCAC 121  
 DB 356 GATCAGCGCTGAAGGTGGCCAAAGCGCGCGGAGGCGAGGGGTAGAGCGCGGCCCT 415  
 OY 122 CATGCAATATCTGTCCTCCGCTTGACAGAGGTCAGACACGATGATCGTCCGGCGCGT 181  
 DB 416 CGTCAACTGTCATGAGCTTCGACAAAGCCACAGCAGCAGGTGCCCAACCGGACGTCAT 475  
 OY 182 CGAGTGCCTGG 192  
 DB 476 CAAGGAGACCG 486

RESULT 12  
 BE917902  
 LOCUS BE917902 550 bp mRNA linear EST 29-SEP-2000  
 DEFINITION OVL\_7\_E07.b1\_A002 Ovary 1 (OVL) Sorghum bicolor cDNA, mRNA  
 sequence.  
 ACCESSION BE917902  
 VERSION BE917902.1 GI:10420470  
 KEYWORDS EST.  
 SOURCE Sorghum.  
 ORGANISM Sorghum bicolor

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC  
 clade; Panicoideae; Andropogoneae; Sorghum.



High quality sequence stop: 552

## FEATURES

POLYA-No. Location/Qualifiers

1..555

/organism="Sorghum bicolor"

/db\_xref="taxon:4558"

/clone\_1lb="Ovary 1 (OVI)"

/note="Organ: Mix of ovaries of varying immature stages from 8-week-old plants: Vector: pBluescript II from Lambda Zap II; Site\_1: XhoI; Site\_2: EcoRI; The library was made from poly-A RNA in the cloning vector lambda Zap II. Clones to be sequenced were prepared by mass excision."

## BASE COUNT

108 a 154 c 179 g 114 t

## ORIGIN

## Query Match

18.8% Score 39; DB 12; Length 555;

Best Local Similarity 50.3%; Pred. No. 2.8;

Matches 96; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

QY 2 CAAGAGAAAGCGCTACGACCTGCGCCGCTGCTGCTATTAAGTCAACAGAGCTCGG 61

Db 94 CAAGGTGTGTGTGTACAGGTGTGTGACGCGGAGCTGTGACTTCTACAAAGACTTCAA 153

QY 62 GCTGCATGTGCGGAGCCCATCATCGTGTGACGCTGACCAAGAGAGCTGTGCGCCAC 121

Db 154 GATCAGCGCTGAAGGTGGCCAGGCGCGCGCGGAGGCGAGGGTGAAGCGCGCGCGCT 213

QY 122 CATCGAATATCTGTGCTGCGCTGACAGAGGTGACAGCAGATGCTGCGGCGCGCT 181

Db 214 CGTCACTGTGCTCATGAGCTGCAAGAGGCAAGCAGCAGGTGCCCGACCGAGCTCAT 273

QY 182 CGAGGTGCCG 192

Db 274 CAAGAGAGACCG 284

## RESULT 15

BG605810

## LOCUS

RH122.81.G04.g1\_A003 RH122) Sorghum prolinquum cDNA, mRNA

## DEFINITION

## ACCESSION

BG605810

## VERSION

BG605810.1

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

BASE COUNT 105 a 157 c 176 g 118 t 1 others

Clones to be sequenced were prepared by mass excision."

ORIGIN

Query Match 18.8% Score 39; DB 12; Length 557;

Best Local Similarity 50.3%; Pred. No. 2.8;

Matches 96; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

QY 2 CAAGAGAAAGCGCTACGACCTGCGCCGCTGCTGCTATTAAGTCAACAGAGCTCGG 61

Db 81 CAAGGTGTGTGTGTACAGGTGTGTGACGCGGAGCTGTGACTTCTACAAAGACTTCAA 140

QY 62 GCTGCATGTGCGGAGCCCATCATCGTGTGACGCTGACCAAGAGAGCTGTGCGCCAC 121

Db 141 GATCAGCGCTGAAGGTGGCCAGGCGCGCGCGGAGGCGAGGGTGAAGCGCGCGCGCT 200

QY 122 CATCGAATATCTGTGCTGCGCTGACAGAGGTGACAGCAGATGCTGCGGCGCGCT 181

Db 201 CGTCACTGTGCTCATGAGTTCGACAAAGGCCAAGCAGCAGGTGCCCGACCGAGCTCAT 260

QY 182 CGAGGTGCCG 192

Db 261 CAAGAGAGACCG 271

Search completed: November 13, 2002, 04:00:49

Job time : 1118.63 secs

FEATURES

source

1..557

Location/Qualifiers

/organism="Sorghum prolinquum"

/db\_xref="taxon:132711"

/clone\_1lb="RH122"

/note="Organ: Rhizomes; Vector: pBluescript II from Lambda Zap II; Site\_1: XhoI; Site\_2: EcoRI; The library was made from poly-A RNA in the cloning vector lambda Zap II."

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using SW model

Run on: November 12, 2002, 16:05:40 ; Search time 21.1607 Seconds  
(Without alignments)  
3487.380 Million cell updates/sec

Title: US-09-697-123b-22

Sequence: 1 tcaaggagaagcgctacgac.....ccgltggaacgcagacat 208

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 320260 seqs, 177392727 residues

Total number of hits satisfying chosen parameters: 640520

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications\_NA:\*

1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/2/pubpna/PCY\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/2/pubpna/PCYUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*  
10: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*  
12: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*  
13: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*  
14: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	88	42.3	5096	10 US-09-984-711-5	Sequence 5, Appl1
2	88	42.3	5099	9 US-10-075-460-5	Sequence 5, Appl1
3	88	42.3	5099	10 US-09-887-052-1	Sequence 1, Appl1
4	88	42.3	5099	10 US-09-887-052-3	Sequence 3, Appl1
5	88	42.3	5099	10 US-09-887-052-5	Sequence 5, Appl1
6	35.6	17.1	3756	10 US-09-841-132-425	Sequence 425, App
7	34	16.3	252	10 US-09-878-574-8901	Sequence 8901, App
8	34	16.3	680	10 US-09-841-132-267	Sequence 267, App
9	32.4	15.6	1437	10 US-09-815-242-7738	Sequence 7738, App
10	32	15.4	711	10 US-09-815-242-7723	Sequence 7723, App
11	31.4	15.1	1164	10 US-09-815-242-7696	Sequence 7696, App
12	31.4	15.1	11220	10 US-09-861-289-32	Sequence 32, Appl1
13	31.4	15.1	36778	10 US-09-861-289-5	Sequence 5, Appl1
14	31.2	15.0	278	10 US-09-294-093B-1392	Sequence 1392, App
15	31.2	15.0	704	12 US-10-062-254-109	Sequence 109, App
16	31.2	15.0	752	12 US-10-062-254-111	Sequence 111, App
17	31.2	15.0	936	10 US-09-815-242-7776	Sequence 7776, App
18	31.2	15.0	8673	10 US-09-815-242-7474	Sequence 7474, App
19	31	14.9	1287	12 US-10-029-715-3	Sequence 3, Appl1

20	31	14.9	3084	10 US-09-881-752A-79	Sequence 79, Appl1
21	31	14.9	8673	10 US-09-815-242-7313	Sequence 7313, App
22	30.8	14.8	337	10 US-09-960-352-10223	Sequence 10223, App
23	30.8	14.8	1293	10 US-09-815-242-7978	Sequence 7978, App
24	30.8	14.8	1296	10 US-09-815-242-4090	Sequence 4090, App
25	30.8	14.8	1440	10 US-09-796-858-7	Sequence 7, Appl1
26	30.8	14.8	1534	10 US-09-765-831A-20	Sequence 20, Appl1
27	30.6	14.7	264	10 US-09-923-876-4989	Sequence 4989, App
28	30.6	14.7	953	10 US-09-452-239-45	Sequence 45, Appl1
29	30.6	14.7	1266	10 US-09-815-242-7920	Sequence 7920, App
30	30.2	14.5	585	10 US-09-870-162A-10	Sequence 10, Appl1
31	30.2	14.5	1298	10 US-09-825-300-682	Sequence 682, App
32	30.2	14.5	2109	10 US-09-815-242-7761	Sequence 7761, App
33	30.2	14.5	6252	10 US-09-964-824A-23	Sequence 23, Appl1
34	30.2	14.5	14272	10 US-09-870-162A-23	Sequence 23, Appl1
35	30	14.4	1455	10 US-09-887-576-782	Sequence 782, App
36	30	14.4	2329	10 US-09-816-828-9	Sequence 9, Appl1
37	30	14.4	4863	9 US-09-815-242-4071	Sequence 4071, App
38	29.8	14.3	592	10 US-09-884-467-6	Sequence 6, Appl1
39	29.8	14.3	1305	10 US-09-815-242-7730	Sequence 7730, App
40	29.8	14.3	2430	10 US-09-861-289-23	Sequence 23, Appl1
41	29.8	14.3	13613	10 US-09-861-289-3	Sequence 3, Appl1
42	29.6	14.2	897	10 US-09-815-242-4001	Sequence 4001, App
43	29.6	14.2	914	10 US-09-964-824A-553	Sequence 553, App
44	29.6	14.2	914	10 US-09-880-107-2357	Sequence 2357, App
45	29.6	14.2	1115	9 US-09-981-353-126	Sequence 126, App

ALIGNMENTS

RESULT 1  
US-09-984-711-5  
; Sequence 5, Application US/09984711  
; Patent No. US20020119549A1  
; GENERAL INFORMATION:  
; APPLICANT: MOECKEL, Bettina  
; APPLICANT: BATHE, Brigitte  
; APPLICANT: STEPHAN, Hans  
; APPLICANT: KREUTZER, Caroline  
; APPLICANT: HERMANN, Thomas  
; APPLICANT: PFEFFERLE, Walter  
; APPLICANT: BINDER, Michael  
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES WHICH CODE FOR THE TP51 GENE  
; FILE REFERENCE: 204209USO  
; CURRENT APPLICATION NUMBER: US/09/984,711  
; PRIOR FILING DATE: 2001-10-31  
; PRIOR APPLICATION NUMBER: DE10108230.9  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 5  
; LENGTH: 5096  
; TYPE: DNA  
; ORGANISM: Corynebacterium glutamicum  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (702)..(4196)  
; OTHER INFORMATION:  
US-09-984-711-5

Query Match 42.3%; Score 88; DB 10; Length 5096;  
Best Local Similarity 68.5%; Pred. No. 3.5e-17;  
Matches 137; Conservative 0; Mismatches 60; Indels 3; Gaps 1;  
QY 9 AAGGCTACACCTGGCGCGTGGCTATTAAGCTCAACAGAGCTGGCTCAT 68  
DB 1578 AAGGCTACACCTGGCGCGTGGCTATTAAGCTCAACAGAGCTGGCTCAT 68  
QY 69 GTGGCGAGCCATACCTGGCTGGAGCTGAGAGAGAGAGAGAGAGAGAGAG 128  
DB 1635 GTGGCGAGCCATGCTGGCTGGAGCTGAGAGAGAGAGAGAGAGAGAGAG 1634



DB 1635 GTGGGACACAGATGTTGATGACTCTTACTGAAGAGACATGCAACCCACCATTCAG 1694  
QY 129 TATCTGTCCTGCTTGCACGAGGTCAACACGATGATGTTCCGGCGGCGTCGAGGTG 188  
DB 1695 TACCTGTGCTGCTGCACGAGGTGAGCGCGTCATGACTTCTCAAAATGGTGAAGAGATC 1754  
QY 189 CCGGTGGAACCGACGACAT 208  
DB 1755 CCAGTCGAGACCGATGACAT 1774

RESULT 5  
US-09-887-052-5

; Sequence 5, Application US/09887052  
; Patent No. US20020119537A1  
; GENERAL INFORMATION:  
; APPLICANT: MOECKEL, Bettina  
; APPLICANT: BATHE, Brigitte  
; APPLICANT: HERMANN, Thomas  
; APPLICANT: PFEFFERLE, Walter  
; APPLICANT: BINDER, Michael  
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES WHICH CODE FOR THE rpoB GENE  
; FILE REFERENCE: 204212050X  
; CURRENT APPLICATION NUMBER: US/09/887,052  
; PRIOR FILING DATE: 2001-06-25  
; PRIOR APPLICATION NUMBER: DE10107229.5  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: Patentin version 3.0  
; SEQ ID NO 5  
; LENGTH: 5099  
; TYPE: DNA  
; ORGANISM: Corynebacterium glutamicum  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (702)..(4196)  
US-09-887-052-5

Query Match 42.3%; Score 88; DB 10; Length 5099;  
Best Local Similarity 68.5%; Pred. No. 3.5e-17;  
Matches 137; Conservative 0; Mismatches 60; Indels 3; Gaps 1;

QY 9 AAGCCCTACGACCTGGCCCGCTGCTGCTATTAAGTCAACAAGAGCTGGGCTCAT 68  
DB 1578 AAGCCCTACGACCTGGCTGCTGCTGCTTACAAAGATCAACCCAGCTCGGCT--T 1634  
QY 69 GTGGGAGCCCATCAGCTCTGTCAGCGTGAACGAGAGAGCTGCGCCACCATCGAA 128  
DB 1635 GTGGGAGCCCATCAGCTCTGTCAGCGTGAACGAGAGAGCTGCGCCACCATCGAG 1694  
QY 129 TATCTGTCCTGCTTGCACGAGGTCAACACGATGATGTTCCGGCGGCGTCGAGGTG 188  
DB 1695 TACCTGTGCTGCTGCACGAGGTGAGCGCGTCATGACTTCTCAAAATGGTGAAGAGATC 1754  
QY 189 CCGGTGGAACCGACGACAT 208  
DB 1755 CCAGTCGAGACCGATGACAT 1774

## RESULT 6

US-09-841-132-425  
; Sequence 425, Application US/09841132  
; Patent No. US20020061848A1  
; GENERAL INFORMATION:  
; APPLICANT: Bhatia, Ajay  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Probst, Peter  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND  
; FILE REFERENCE: 210121,469C8  
; CURRENT APPLICATION NUMBER: US/09/841,132  
; CURRENT FILING DATE: 2001-04-23

; NUMBER OF SEQ ID NOS: 599  
; SOFTWARE: FastSeq for Windows Version 3.0/4.0  
; SEQ ID NO 425  
; LENGTH: 3756  
; TYPE: DNA  
; ORGANISM: Chlamydia trachomatis serovar D  
US-09-841-132-425

Query Match 17.1%; Score 35.6; DB 10; Length 3756;  
Best Local Similarity 56.5%; Pred. No. 0.035;  
Matches 87; Conservative 0; Mismatches 64; Indels 3; Gaps 1;

QY 1 TCAGAGGAGAGCGCTACGACCTGGCCCGCTGCTGCTATAGGTCAACAAGAGCTCG 60  
DB 983 TCGATGCTAAAGCGTATATATTAGCCCGCTTGGACGTTATTAATTAATAAATTAG 1042  
QY 61 GCGTGC---ATGTGGCGAGCCCATCAGCTGTCGACCGCTTACCGAAGACGCTGTG 117  
DB 1043 GATTCCTATTAGACGAGCAAAATATCTCAAGTCACTTGGAAAAAGAGATGTATCG 1102  
QY 118 CCACATCGAATATCTGTCGCTGCTGACGAGG 151  
DB 1103 GCGCGTTGAATATTTGATTCGTTGCAATGGC 1136

RESULT 7  
US-09-878-574-8901

; Sequence 8901, Application US/09878574  
; Patent No. US20020110548A1  
; GENERAL INFORMATION:  
; APPLICANT: Byrum, Joseph R.  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Thompson, Michael D.  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; FILE REFERENCE: 38-21(15401)B  
; CURRENT APPLICATION NUMBER: US/09/878,574  
; PRIOR FILING DATE: 2001-12-21  
; PRIOR APPLICATION NUMBER: 09/333,535  
; NUMBER OF SEQ ID NOS: 15775  
; SEQ ID NO 8901  
; LENGTH: 252  
; TYPE: DNA  
; ORGANISM: Glycine max  
; OTHER INFORMATION: Clone ID: 701101828H1  
US-09-878-574-8901

Query Match 16.3%; Score 34; DB 10; Length 252;  
Best Local Similarity 50.0%; Pred. No. 0.16;  
Matches 85; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

QY 17 GACCTGGCCCGCTGCTGCTGCTATTAAGTCAACAAGAGCTGGGCTGATGTCGGCA 76  
DB 82 GCGCGTGGCCCGCATGAGCGACCCGACGCTATCAAGGCCATACGCGGCGCTCACAT 141  
QY 77 GCCCATCAGCTGTCGACGCTGACCGAAGAGAGCTGTCGCGCCACCATCAATATCTGT 136  
DB 142 CCCGTCATGCGCCACGCGCCGACATCGGCGCACCTTCGTGAGGCGCCACATCTCGAGGCCCT 201  
QY 137 CCGCTGACGAGGGTCAAGCACGATGATGCTTCGGGCGGCGTGTGAGC 186  
DB 202 CCGCATGACATCTGACGACGCGCGGAGGTCCTACCCCTCCGACGAGC 251

## RESULT 8

US-09-841-132-267  
; Sequence 267, Application US/09841132  
; Patent No. US20020061848A1  
; GENERAL INFORMATION:  
; APPLICANT: Bhatia, Ajay  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Probst, Peter

;; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND  
;; FILE REFERENCE: 210121.469c8  
;; CURRENT APPLICATION NUMBER: US/09/841,132  
;; NUMBER OF SEQ ID NOS: 599  
;; SOFTWARE: FastSeq for Windows Version 3.0/4.0  
;; SEQ ID NO 267  
;; LENGTH: 680  
;; TYPE: DNA  
;; ORGANISM: Chlamydia  
US-09-841-132-267

Query Match 16.3%; Score 34; DB 10; Length 680;  
Best Local Similarity 55.8%; Pred. No. 0.2;  
Matches 86; Conservative 0; Mismatches 65; Indels 3; Gaps 1;

OY 1 TCAAGAGAGCCGCTACGACCTGGCCGCTGCTGCTATAGTCAAGAGAGCTCG 60  
DB 442 TCGATCGTAAACGTTAAATTTAGCGCGCTTGACGTTATAAATTAATAAATTG 501  
OY 61 GGTGCG--ATGTCGGCGAGCCCATACGTCGTCGACCGTGCAGGAGAGCGTGTG 117  
DB 502 GCTTCCCATTTAGAGAGCAACCTTATCTCAAGTACTTTGAGAAAAAGAGATTATCG 561  
OY 118 CCACCATCGAATATCTGTCGCTCGCTTCGACGAGGG 151  
DB 562 GCCGCTTGAATAATTTGATTGCTTTCGAATGGG 595

RESULT 9  
US-09-815-242-7738  
; Sequence 7738, Application US/09815242  
; Patent No. US20020061569A1  
; GENERAL INFORMATION:  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Karl L.  
; APPLICANT: Zyskind, Judith W.  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John D.  
; APPLICANT: Carr, Grant J.  
; APPLICANT: Yamamoto, Robert T.  
; APPLICANT: Xu, H. Howard  
; TITLE OF INVENTION: Identification of Essential Genes in  
; FILE REFERENCE: ELITRA.011A  
; CURRENT APPLICATION NUMBER: US/09/815,242  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; NUMBER OF SEQ ID NOS: 14110  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7738  
; LENGTH: 1437  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)...(1437)  
US-09-815-242-7738

Query Match 15.6%; Score 32.4; DB 10; Length 1437;  
Best Local Similarity 52.2%; Pred. No. 0.69;  
Matches 72; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

OY 46 TCACACAGAGCTGGCGGTGCATGTCGGCGAGCCCATCAGCTGTCGACGCTGACCGAG 105  
DB 809 TCGACAGCTGATCGTCCGCTCGCCGCTCGCCGAGACCCAGACCTGCGCCCGC 868  
OY 106 AAGACGTCGTGCGCCACATCGAATATCTGCGCTTGACAGAGGTGACACAGATGA 165  
DB 869 ACAGCGGCGTGACCCCTGGACAGCGGTGCTTACATCTAGCTGACAGACACTGCAAGACA 928  
OY 166 TCGTTCGGCGCGCTCG 183  
DB 929 GCGTTCGCGCGCTCTCG 946

RESULT 10  
US-09-815-242-7723  
; Sequence 7723, Application US/09815242  
; Patent No. US20020061569A1  
; GENERAL INFORMATION:  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Karl L.  
; APPLICANT: Zyskind, Judith W.  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John D.  
; APPLICANT: Carr, Grant J.  
; APPLICANT: Yamamoto, Robert T.  
; APPLICANT: Xu, H. Howard  
; TITLE OF INVENTION: Identification of Essential Genes in  
; FILE REFERENCE: ELITRA.011A  
; CURRENT APPLICATION NUMBER: US/09/815,242  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; NUMBER OF SEQ ID NOS: 14110  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7723  
; LENGTH: 711  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)...(711)  
US-09-815-242-7723

Query Match 15.4%; Score 32; DB 10; Length 711;  
Best Local Similarity 55.4%; Pred. No. 0.78;  
Matches 62; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

OY 92 GAGCGTACCCGAGAGAGCTGTCGGCCACCATCATATATGTCGCTGACGAGGG 151  
DB 171 GATGCTGTCGCGGAGAGAGCGCTGTCATCTGCCCCAGTGCAGCCGACG 230  
OY 152 TCAGACCATCATGATGCTTCCGGGCGGTCGAGTGGCCGCTGGAACGAC 203  
DB 231 TCGATCTGATGCTACCGCGCGCACCCAGATATGAGAGAGTACTCGC 282



```

? TITLE OF INVENTION: DNA encoding melhymycin and pikromycin
? FILE REFERENCE: 600.438US1
? CURRENT APPLICATION NUMBER: US/09/861,289
? CURRENT FILING DATE: 2001-05-18
? PRIOR APPLICATION NUMBER: 09/7105,537
? PRIOR FILING DATE: 1998-06-26
? NUMBER OF SEQ. ID NOS: 43
? SOFTWARE: FastSeq for Windows Version 3.0
? SEQ ID NO 32
? LENGTH: 11220
? TYPE: DNA
? ORGANISM: Streptomyces venezuelae
US-09-861-289-32

Query Match          15.1%; Score 31.4; DB 10; Length 11220;
Best local similarity 48.1%; Pred. No. 2.1;
Matches 89; Conservative 0; Mismatches 96; Total 0; Gaps 3

```

Query Match	15.18;	Score 31.4;	DB 10;	Length 11220;
Best Local Similarity	48.18;	Pred. No. 2.1;		
Matches	89;	Conservative	0;	Mismatches 96; Indels 0; Gaps 0;
QY	12	CGTAGCAGCTGGCCCGCGTGGTGGCTTAAGTGCACAAAGAAAGCTGGGGCTGCATGTC	71	
DB	5535	CTCGCAGCAGTCCACGTCGGCGGTCTCTCAGTCGGGGCTTCCGAGCCGCGCCGATGAC	5476	
QY	72	GGCGAGCCCATCAGTGTGTGAGCGCTGACCCGAAGAAGCGTGGGCCACCATCGAATAT	131	
DB	5475	GGCGTGTGGCGAAGGCGCCGTTCCGGGCGGTGACGGCGCTTGCTCCGCGCGTCTTGTTGAC	5416	
QY	132	CTGTGTCGGCTTTCACGAGGAGGTGCACACAGATGATGTTCCGGGGCGGCGTTCAGAGTCCG	191	
DB	5415	CGCGCTCGCGCGCACACGCCGCCACGACACAGCGTGTCCGTGGGCGGGCGTCCACAGGCG	5356	
QY	192	GTGCA	196	
DB	5355	CTGCA	5351	

```

RESULT 13
US-09-861-289-5/C
: Sequence 5, Application US/09861289
: Patent No. US20020110897A1
: GENERAL INFORMATION:
: APPLICANT: Sherman, D.H.
: APPLICANT: Liu, H.
: APPLICANT: Xue, Y.
: APPLICANT: Zhao, L.
: TITLE OF INVENTION: DNA encoding methymycin and pikromycin
: FILE REFERENCE: 600.438US1
: CURRENT APPLICATION NUMBER: US/09/861,289
: CURRENT FILING DATE: 2001-05-18
: PRIOR APPLICATION NUMBER: 09/105,537
: PRIOR FILING DATE: 1998-06-26
: NUMBER OF SEQ. ID NOS: 43
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ. ID NO 5
: LENGTH: 36778
: TYPE: DNA
: ORGANISM: Streptomyces venezuelae
US-09-861-289-5

Query Match          15.1%; Score 31.4; DB 10; Length 36778;
Best Local Similarity 48.1%; Pred. No. 2.6;
Matches 89; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

QY      12  CGGTACGACCTGGCCCGCGTCGTCGGCATATAGTCAACAAAGACTCGGCGTCATGTC  71
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 21222 CTCGACGAGAGTCACAGTCGGCGGTGTCACATCGGCGCTCCGACGAGCGCCGCGGATGAC 21163

QY      72  GGGAGACCATATACGTCTGCACACGCTGACCGAAGAGAGACGTCGTGGCCACATCGAATAT  131
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 21162 GCGCTTCCTCCGAGGCGCCGTTCCGGGGCGGTACAGCCGTTCTCGCGCCGCTCGGTTGAC 21103

QY      132 CTGCTCCGCTTGCACGAGGATCGAACACATATATGTTCCGGCGGCGCTCGAGTGCCG  191
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 21102 CGCGTCGCGCGCCACGACCGCCACACACAGGTGTCGTTGCGGCGGGGCTTCGCACAGGCG 21043

```



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 12, 2002, 15:56:36 : Search time 27.8936 Seconds

(without alignments) 2286.857 Million cell updates/sec

Title: US-09-697-123B-22

Perfect score: 208  
Sequence: 1 tcaagagaagcgtacgac.....ccgtggaaccgcagacat 208

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_NA.\*  
1: /cgn2\_6/ptodata/1/ina/5A.COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B.COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A.COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B.COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PCUS.COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfilest.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	206.4	99.2	970	1 US-08-250-030-1	Sequence 1, Appl
2	206.4	99.2	970	5 PCT-US95-06790-1	Sequence 1, Appl
3	147.2	70.8	3447	2 US-08-313-185-57	Sequence 57, Appl
4	147.2	70.8	3447	3 US-08-082-614A-57	Sequence 57, Appl
5	38.8	18.7	1412	1 US-08-097-831-1	Sequence 1, Appl
6	37	17.8	6085	4 US-09-029-603-4	Sequence 4, Appl
7	36	17.3	1894	4 US-09-329-350-12	Sequence 32, Appl
8	34.6	16.6	1621	1 US-08-722-001-13	Sequence 13, Appl
9	34.6	16.6	1776	1 US-08-722-001-29	Sequence 29, Appl
10	34.6	16.6	2140	1 US-08-314-698-1	Sequence 1, Appl
11	34.6	16.6	2140	1 US-08-228-932-1	Sequence 1, Appl
12	34.6	16.6	2140	1 US-08-468-939-1	Sequence 1, Appl
13	34.6	16.6	2140	2 US-08-406-855A-1	Sequence 1, Appl
14	34.6	16.6	2140	2 US-08-722-190-1	Sequence 1, Appl
15	34.6	16.6	2140	3 US-08-244-354-1	Sequence 1, Appl
16	34.6	16.6	2140	3 US-09-206-899-1	Sequence 1, Appl
17	34.6	16.6	2140	4 US-09-444-783-1	Sequence 1, Appl
18	34.6	16.6	2140	4 US-09-688-415-1	Sequence 1, Appl
19	34.6	16.6	2140	5 PCT-US95-04203-1	Sequence 1, Appl
20	34.2	16.4	22306	4 US-09-453-702B-251	Sequence 251, App
21	34	16.3	680	4 US-09-556-877-267	Sequence 267, App
22	34	16.3	680	4 US-09-620-412C-267	Sequence 267, App
23	32.8	15.8	30001	1 US-08-125-468-1	Sequence 1, Appl
24	32.8	15.8	30001	2 US-08-474-933-1	Sequence 1, Appl
25	32.6	15.7	734	4 US-09-221-017B-1070	Sequence 1070, Ap
26	32.4	15.6	38584	4 US-09-453-702B-50	Sequence 50, Appl
27	32.4	15.6	48908	4 US-09-453-702B-137	Sequence 137, App

28	32	15.4	1146	3 US-08-911-853-26	Sequence 26, Appl
29	32	15.4	1146	4 US-09-479-409-26	Sequence 26, Appl
30	32	15.4	1146	4 US-09-479-453-26	Sequence 26, Appl
31	32	15.4	17612	3 US-08-911-853-29	Sequence 29, Appl
32	32	15.4	17612	4 US-09-479-409-29	Sequence 29, Appl
33	32	15.4	17612	4 US-09-479-453-29	Sequence 29, Appl
34	31.8	15.3	2122	4 US-09-029-603-1	Sequence 1, Appl
35	31.8	15.3	2639	1 US-07-952-817-8	Sequence 8, Appl
36	31.8	15.3	2639	6 5210025-1	Patent No. 5210025
37	31.4	15.1	1268	3 US-08-713-569-5	Sequence 5, Appl
38	31.4	15.1	1320	2 US-08-461-775-8	Sequence 8, Appl
39	31.4	15.1	1320	3 US-08-031-606-8	Sequence 8, Appl
40	31.4	15.1	1620	2 US-08-461-775-10	Sequence 10, Appl
41	31.4	15.1	1620	3 US-09-031-606-10	Sequence 10, Appl
42	31.4	15.1	2167	2 US-08-461-775-9	Sequence 9, Appl
43	31.4	15.1	2167	3 US-09-031-606-9	Sequence 9, Appl
44	31.4	15.1	2668	2 US-08-461-775-11	Sequence 11, Appl
45	31.4	15.1	2668	3 US-09-031-606-11	Sequence 11, Appl

## ALIGNMENTS

```
RESULT 1
US-08-250-030-1
: Sequence 1, Application US/08250030
: Patent No. 5643723
:
: APPLICANT: Persing, David H.
: TITLE OF INVENTION: Detection of a Genetic Locus Encoding
: TITLE OF INVENTION: Resistance to Rifampin in Mycobacterial Cultures and in
: TITLE OF INVENTION: Clinical Specimens
: NUMBER OF SEQUENCES: 15
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Schwegman, Lundberg & Woessner
: STREET: 3500 IDS Center
: CITY: Minneapolis
: STATE: MN
: COUNTRY: USA
: ZIP: 55402
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/250,030
: FILING DATE: 26-MAY-1994
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Mueiting, Ann M.
: REGISTRATION NUMBER: 33,977
: REFERENCE/DOCKET NUMBER: 150.1050S1
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 612-339-0331
: TELEFAX: 612-339-3061
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 970 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA
: US-08-250-030-1

Query Match          99.2%: Score 206.4: DB 1: Length 970:
Best Local Similarity 99.5%: Pred. No. 4.5e+48:
Matches 207: Conservative 0: Mismatches 1: Indels 0: Gaps 0:

DB 1 TCAAGAGAAGCGCTACGACCTGGCCCGGCTGCTATTAAGCTCAACAAGAGCTCG 60
|||||
26 TCAAGAGAAGCGCTACGACCTGGCCCGGCTGCTATTAAGCTCAACAAGAGCTCG 85
```



Sequence 57, Application US/09082614A  
Patent No. 6124098  
GENERAL INFORMATION:  
APPLICANT: Heyn, Beate  
APPLICANT: Cole, Stewart  
APPLICANT: Young, Douglas  
APPLICANT: Zhang, Ying  
APPLICANT: Honore, Nadine  
APPLICANT: Telenti, Amalio  
APPLICANT: Bodmer, Thomas  
TITLE OF INVENTION: Rapid Detection of Antibiotic Resistance  
TITLE OF INVENTION: In Mycobacterium Tuberculosis  
NUMBER OF SEQUENCES: 66  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &  
ADDRESS: Dinner  
STREET: 1300 I Street, N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20005-3315  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/082,614A  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/313,185  
FILING DATE: 12-OCT-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Meyers, Kenneth J.  
REGISTRATION NUMBER: 25,146  
REFERENCE/DOCKET NUMBER: 02356, 0068-00000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 408-4400  
TELEFAX: (202) 408-4400  
INFORMATION FOR SEQ ID NO: 57:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3447 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-09-082-614A-57

Query Match 70.8%; Score 147.2; DB 3; Length 3447;  
Best Local Similarity 81.7%; Pred. No. 9,2e-32;  
Matches 170; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY 1 TCAGAGAAAGCGCTACGACCTGGCCGCTGCTGCTATTAAGTCAACAAGAACTCG 60  
DB 809 TCAGAGAAAGCGCTACGACCTGGCCGCTGCTGCTTACAAAGTCAACAAGAACTCG 868  
QY 61 GCGTCGATGTCGGCGAGCCATCAGTCGTCGACGCTACCGAAGAAAGTCTGGCGCA 120  
DB 869 GCGTCGACCGCGGTGATGATCAGTCGTCGACGCTACCGAAGAAAGATGTCGTCGCA 928  
QY 121 CCATCGAATATCTGCTGCGCTTGCACAGAGGTCACAGCAGCATGATCTTCGGGCGGCG 180  
DB 929 CCATAGAGTACCTGCTGCTGTCATGAGAGGTCATGACAAATGACGTCTCCAGTGGGG 988  
QY 181 TCGAGGTGCGGTGGAACCGCAGCAT 208  
DB 989 TAGAGTGCACGTGGAACCTGACATAT 1016

RESULT 5  
US-08-097-831-1  
Sequence 1, Application US/08097831

Patent No. 5510473  
GENERAL INFORMATION:  
APPLICANT: Camerini-Otero, Rafael D.  
APPLICANT: Angov, Evangelina  
TITLE OF INVENTION: Cloning and Expression of Taq recA  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: KNOBBE, MARTENS, OLSON AND BEAR  
STREET: 620 Newport Center Drive 16th Floor  
CITY: Newport Beach  
STATE: CA  
COUNTRY: USA  
ZIP: 92660  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/097,831  
FILING DATE: 19930726  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Altman, Daniel E.  
REGISTRATION NUMBER: 34,115  
REFERENCE/DOCKET NUMBER: NIH066,001A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 714-760-0404  
TELEFAX: 714-760-9502  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1412 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
IMMEDIATE SOURCE:  
CLONE: Taq recA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 62..1082  
US-08-097-831-1

Query Match 18.7%; Score 38.8; DB 1; Length 1412;  
Best Local Similarity 54.1%; Pred. No. 0.039;  
Matches 79; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

QY 50 CAGAAGCTCGGCGTCATGTCGGGAGCCCATCATCAGTCGTCGACGCTGACCGAAGAAGA 109  
DB 370 CAGAAGCTCGGCGTCGAGAGTCGACAGACCTTCTGCTCCAGCCGAGACACCGGGAGCA 429  
QY 110 CGTCGTGCGCACCATCGAATATCTGTCGCTGTCAGCAGAGGTGACAGCAGATGATCGT 169  
DB 430 GCGTTTGAGATCTGTCGAGACTTCTGCGCCGCTCGGGGCGGTGACGATGATCGTGTGA 489  
QY 170 TCGGGGCGCGTCGAGGTGCCGCTG 195  
DB 490 TTCGTGTCGCGCTTGTGTGCCAAG 515

RESULT 6  
US-09-029-603-4  
Sequence 4, Application US/09029603  
Patent No. 6210935  
GENERAL INFORMATION:  
APPLICANT: Schupp, Thomas  
APPLICANT: Engel, Natalie  
APPLICANT: Bietenader, Jurg  
APPLICANT: Toupet, Christine  
APPLICANT: Pospiech, Andreas  
TITLE OF INVENTION: Staurosporin Biosynthesis Gene Clusters

COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/329,350  
FILING DATE: Herewith  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/841,636  
FILING DATE: 30-APR-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/005,335  
FILING DATE: 17-OCT-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/007,926  
FILING DATE: 04-DEC-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/020,840  
FILING DATE: 28-JUN-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/732,181  
FILING DATE: 16-OCT-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/FI96/00550  
FILING DATE: 17-OCT-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Shea Jr., Timothy  
REGISTRATION NUMBER: 41,306  
REFERENCE/DOCKET NUMBER: 1716.051006/MAC/TJS  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)371-2800  
TELEFAX: (202)371-2540  
INFORMATION FOR SEQ ID NO: 32:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1894 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
ORIGINAL SOURCE:  
ORGANISM: Melanocarpus albomyces  
STRAIN: ALK04237  
FEATURE:  
NAME/KEY: exon  
LOCATION: 233..838  
OTHER INFORMATION: /product= "50K-cellulase"  
NAME/KEY: exon  
LOCATION: 916..1596  
OTHER INFORMATION: /product= "50K-cellulase"  
US-09-329-350-32

Query Match 17.3%; Score 36; DB 4; Length 1894;  
Best Local Similarity 56.9%; Pred. No. 0.24; Indels 0; Gaps 0;  
Matches 66; Conservative 0; Mismatches 50;

DB 1152 GCCCTTACCGTGTCGACGCTGACCGAAGAACGTCTGTGGCCACCATCGAATATCTGTT 136  
          |||||     |||     |||     |||     |||     |||     |||     |||     |||  
QY 137 CCGCTGCACGAGGCTCAACACCATGATGCTTCCGGGCGGCTCGAGGTGCCGG 192  
          |||     |||     |||     |||     |||     |||     |||     |||     |||  
Db 1212 CCGGCTGTACGTGCAGACGACGAAGTATGTAAGTCTGTGACGCGCCGG 1267

RESULT 8  
US-08-722-001-13  
Sequence 13, Application US/08722001  
Patent No. 5760054

```

: GENERAL INFORMATION:
: APPLICANT: Thompson, Wayne J.
: APPLICANT: Huff, Joel R.
: APPLICANT: Nerenberg, Jennie B.
: APPLICANT: Lee, Hee-Yoon
: APPLICANT: Bell, Ian M.
: TITLE OF INVENTION: ALPHA1C ADRENERGIC RECEPTOR ANTAGONISTS
: NUMBER OF SEQUENCES: 35
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Merck & Co., Inc.
: STREET: 126 Lincoln Avenue
: CITY: Rahway
: STATE: New Jersey
: COUNTRY: United States of America
: ZIP: 07065
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/722,001
: FILING DATE:
: CLASSIFICATION: 514
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/229,276
: FILING DATE: 14-APR-1995
: ATTORNEY/AGENT INFORMATION:
: NAME: Appollina, Mary A.
: REGISTRATION NUMBER: 34,087
: REFERENCE/DOCKET NUMBER: 19169Y
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (908)594-3462
: TELEFAX: (908)594-4720
: TELEX: 138825
: INFORMATION FOR SEQ. ID NO.: 13:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1621 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: both
: MOLECULE TYPE: cDNA
: HYPOTHEICAL: NO
: ANTI-SENSE: NO
: US-08-722-001-13

Query Match          16.6%; Score 34.6; DB 1; Length 1621;
Best Local Similarity 50.3%; Pred. No. 0.57;
Matches 85; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

QY 40 ATAAAGTCAACAAGAGCTCGGGCTGCATGTCGGCGAGCCCATCACTGTCGACGCTGA 99
DB 312 AGACCGTCAACCACTATTTCATGCTGAACCTGCGCGTGGCGACCTGCTGAGCGCCA 371
QY 100 CCGAAGAAGACGTGCGTGGCCACCATGCAATATCTGTCGCGCTTGACGAGGCTGACGACCA 159
DB 372 CCGTACTGCCCTTCTGCGCCACCATGAGAGTTCTGAGGCTTCTGGGCTTTTGCGCGGCT 431
QY 160 CGATGATCGTTCGGGGCGGCGTGAAGTCCGGTGAAGAAACCGACGACAT 208
DB 432 TCTGCGACGTATGCGCCCGCGTGACGTGCTGCTGCACGCGCTTCAT 480

RESULT 9
US-08-722-001-29
: Sequence 29, Application US/08722001
: Patent No. 5760054
: GENERAL INFORMATION:
: APPLICANT: Thompson, Wayne J.
: APPLICANT: Huff, Joel R.
: APPLICANT: Nerenberg, Jennie B.
: APPLICANT: Lee, Hee-Yoon
: APPLICANT: Bell, Ian M.
```

```

: TITLE OF INVENTION: ALPHA1C ADRENERGIC RECEPTOR ANTAGONISTS
: NUMBER OF SEQUENCES: 35
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Merck & Co., Inc.
: STREET: 126 Lincoln Avenue
: CITY: Rahway
: STATE: New Jersey
: COUNTRY: United States of America
: ZIP: 07065
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/722,001
: FILING DATE:
: CLASSIFICATION: 514
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/229,276
: FILING DATE: 14-APR-1995
: ATTORNEY/AGENT INFORMATION:
: NAME: Appollina, Mary A.
: REGISTRATION NUMBER: 34,087
: REFERENCE/DOCKET NUMBER: 19169Y
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (908)594-3462
: TELEFAX: (908)594-4720
: TELEX: 138825
: INFORMATION FOR SEQ. ID NO.: 29:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1776 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: both
: TOPOLOGY: both
: MOLECULE TYPE: cDNA
: HYPOTHEICAL: NO
: ANTI-SENSE: NO
: US-08-722-001-29

Query Match          16.6%; Score 34.6; DB 1; Length 1776;
Best Local Similarity 50.3%; Pred. No. 0.58;
Matches 85; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

QY 40 ATAAAGTCAACAAGAGCTCGGGCTGCATGTCGGCGAGCCCATCACTGTCGACGCTGA 99
DB 440 AGACCGTCAACCACTATTTCATGCTGAACCTGCGCGTGGCGACCTGCTGAGCGCCA 499
QY 100 CCGAAGAAGACGTGCGTGGCCACCATGCAATATCTGTCGCGCTTGACGAGGCTGACGACCA 159
DB 500 CCGTACTGCCCTTCTGCGCCACCATGAGAGTTCTGAGGCTTCTGGGCTTTTGCGCGGCT 559
QY 160 CGATGATCGTTCGGGGCGGCGTGAAGTCCCGGTGAAGAAACCGACGACAT 208
DB 560 TCTGCGACGTATGCGCCCGCGTGACGTGCTGCTGCACGCGCTTCAT 608

RESULT 10
US-08-334-698-1
: Sequence 1, Application US/08334698
: Patent No. 5556753
: GENERAL INFORMATION:
: APPLICANT: Jonathan A. Bard et al.
: TITLE OF INVENTION: DNA Encoding Human Alpha 1 Adrenergic
: TITLE OF INVENTION: Receptors and Uses Thereof
: NUMBER OF SEQUENCES: 6
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: COOPER & DUNHAM
: STREET: 30 Rockefeller Plaza
: CITY: New York
: STATE: New York
: COUNTRY: U.S.A.
: ZIP: 10112
```

```

: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.24
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/334,698
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US/07/952,798
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: White, John P.
: REGISTRATION NUMBER: 28,678
: REFERENCE/DOCKET NUMBER: 376901
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (212) 977-9550
: TELEFAX: (212) 664-0525
: TELEX: (212) 422523 COOP UI
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2140 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: unknown
: MOLECULE TYPE: DNA (genomic)
: HYPOTHEICAL: N
: ANTI-SENSE: N
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 178..1893
: OTHER INFORMATION:
:
: US-08-334-698-1
:
: Query Match 16.6%; Score 34.6; DB 1; Length 2140:
: Best Local Similarity 50.3%; Pred. No. 0.6;
: Matches 85; Conservative 0; Mismatches 84; Indels 0; Gaps 0:
:
: Oy 40 ATAAGTCAACAGAAAGCTGGGCTGCATGTCGCGAGCCCATCAGTCGTCGAGCTGA 99
: Db 560 AGACCGTCACCACTATTTCATGTAACCTGGCGGCCGCTGCGTCGAGCGCCA 619
: Oy 100 CCGAAGAAAGACGTCGCGCCACCATGCAATATCTGTCGCTTGCACGAGGCTCAGACCA 159
: Db 620 CCGTACTGCCCTTCTCGGCCACCATGAGGTTCTGGGCTTGGGCCCTTGGCCGCCCT 679
: Oy 160 CGATGATCGTTCCGCGCGGCTGAGTGCCGCTGGAACCGACGACAT 208
: Db 680 TCTCGACGTATGGCCCGCGCTGAGCTGCTGTCTGCACGCGCTTCAT 728
:
: RESULT 11
: US-08-228-932-1
: Sequence 1, Application US/08228932
: Patent No. 5578611
: GENERAL INFORMATION:
: APPLICANT: Charles Gluchowski, Carlos C. Ferray, George Chiu,
: APPLICANT: Theresa A. Branchek, John M. Weitzel and Paul R. Hartig
: TITLE OF INVENTION: USE OF ALPHA-1C SPECIFIC COMPOUNDS TO TREAT BENIGN
: TITLE OF INVENTION: PROSTATIC HYPERPLASIA
: NUMBER OF SEQUENCES: 6
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: COOPER & DUNHAM
: STREET: 30 Rockefeller Plaza
: CITY: New York
: STATE: New York
: COUNTRY: U.S.A.
: ZIP: 10112
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.24
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/468,939
: FILING DATE:
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: White, John P.
: REGISTRATION NUMBER: 28,678
: REFERENCE/DOCKET NUMBER: 41878-B/JPM/TEP
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (212) 977-9550
: TELEFAX: (212) 664-0525
: TELEX: (212) 422523 COOP UI
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2140 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: unknown
: MOLECULE TYPE: DNA (genomic)
: HYPOTHEICAL: N
: ANTI-SENSE: N
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 178..1893
: OTHER INFORMATION:
:
: US-08-228-932-1
:
: Query Match 16.6%; Score 34.6; DB 1; Length 2140:
: Best Local Similarity 50.3%; Pred. No. 0.6;
: Matches 85; Conservative 0; Mismatches 84; Indels 0; Gaps 0:
:
: Oy 40 ATAAGTCAACAGAAAGCTGGGCTGCATGTCGCGAGCCCATCAGTCGTCGAGCTGA 99
: Db 560 AGACCGTCACCACTATTTCATGTAACCTGGCGGCCGCTGCGTCGAGCGCCA 619
: Oy 100 CCGAAGAAAGACGTCGCGCCACCATGCAATATCTGTCGCTTGCACGAGGCTCAGACCA 159
: Db 620 CCGTACTGCCCTTCTCGGCCACCATGAGGTTCTGGGCTTGGGCCCTTGGCCGCCCT 679
: Oy 160 CGATGATCGTTCCGCGCGGCTGAGTGCCGCTGGAACCGACGACAT 208
: Db 680 TCTCGACGTATGGCCCGCGCTGAGCTGCTGTCTGCACGCGCTTCAT 728
:
: RESULT 12
: US-08-468-939-1
: Sequence 1, Application US/08468939
: Patent No. 5714381
: GENERAL INFORMATION:
: APPLICANT: Jonathan A. Bard et al.
: APPLICANT: DNA Encoding Human Alpha 1 Adrenergic
: TITLE OF INVENTION: Receptors and Uses Thereof
: TITLE OF INVENTION: Receptors and Uses Thereof
: NUMBER OF SEQUENCES: 6
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: COOPER & DUNHAM LLP
: STREET: 1185 Avenue of the Americas
: CITY: New York
: STATE: New York
: COUNTRY: U.S.A.
: ZIP: 10036
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.24
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/468,939
: FILING DATE:
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: White, John P.
: REGISTRATION NUMBER: 28,678
```





OTHER INFORMATION:  
US-08-722-190-1

Query Match	16.6%;	Score	34.6;	DB	2;	Length	2140;
Best Local Similarity	50.3%;	Pred. No.	0.6;				
Matches	85;	Conservative	0;	Mismatches	84;	Indels	0;
						Gaps	0;

QY 40 ATAAAGTTCACAAAGAGCTCGGGTGCATGTGGGAGGCCCATCACTGTGTCAACGCTGA 99  
 Db 560 AGACGCTCACCAACTATTTCATCGTGAACCTGGCCGTGGCCACACCTGTGTGTGACGCCCA 619  
 QY 100 CCGAAGAAAGACGTCTGGGCCACATCAATATATGTGTCGCTTGGACAGAGGGTAAAGCA 159  
 Db 620 CCGTACGTGCCCTTCTCGGCCACCACTGAGAGTTTGGGCTTCTGGGACCTTTTGGCCGCCGT 679  
 QY 160 CGATGATCGTTCCGGCGGGCGCTCGAGTGGCCGTGGAACACGACGACAT 208  
 Db 680 TCTGCAGAGTATGGGCGCGCGTGGAGCTGTGTGTCTGCACAGGCTTCAT 728

RESULT 15  
US-08-344

US-08-244-354-1  
Sequence 1, Application US/08244354  
Patent No. 6015818

```

: GENERAL INFORMATION:
:
: APPLICANT: Charles Gluchowski, et al.
:
: TITLE OF INVENTION: USE OF ALPHA-1C SPECIFIC COMPOUNDS TO
:
: TITLE OF INVENTION: TREAT BENIGN PROSTATIC HYPERPLASIA
:
: NUMBER OF SEQUENCES: 6
:
: CORRESPONDENCE ADDRESS:

```

COMPUTER READABLE FORM:

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC

```

OPERATING SYSTEM: IBM PC

SOFTWARE: PatentIn Release #1.24

```

; CURRENT APPLICATION DATA:
;
; APPLICATION NUMBER: FIS/08/2344 3E

```

FILING DATE: April 1, 1997

CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:  
NAME: Whitcomb

NAME: WILCE, JOHN F.  
REGISTRATION NUMBER: 28.678

REFERENCE/DOCKET NUMBER: 41878-D-

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 40131, 270 0400

TELEPHONE: (212) 218-0400  
TELEFAX: (212) 391-0525

TELEX:

```

; INFORMATION FOR SEQ ID NO: 1:
SEQUENCE QUALITY

```

SEQUENCE CHARACTERISTICS:  
LENGTH: 2140 base pairs

TYPE: nucleic acid

```
STRANDEDNESS: single
```

MOLECULE TYPE: DNA (genomic)

HYPOTHETICAL: N

ANTI-SENSE: N

NAME/KEY: CDS

LOCATION: 178..1893

OTHER INFORMATION:

05-08-244-33A-1

Query Match	16.68; Score 3
-------------	----------------

Best Local Similarity 50.38; Pred. N

malchies 85; conservative 0; m1sm

40 ATAGGTCACACAAGAAGCTCGGGCTGCATGT

Accession	Sequence	Position
Db	AGACCGCTACCAACTATTTCATCGTGAACCGCGCGCGGCGACCTGGCTGTGAACGCCA	6139
Oy	CCGAAGAAGACGTCGCGGCGCACCATCGAATATCTGTCGCGTTGCACGAGGTCAGACCA	1599
Db	CCGACTGCCCTTCTCGGCGACCATGAGAGGCTCTGGCTTCTCGGCGCTTTGGCGCGCGCT	6799
Oy	CGATGATCGTCCGGGCGGCGTCGAGGCGCGGGGAAACCGACGACGAT	208
Db	TCGCGCAGCATGGGCGCCCGCTGGACGCTGCTGTCTGCACAGGCTTCGAT	728

Search completed: November 12, 2002, 21:03:40  
Job time : 39.8936 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 12, 2002, 15:44:35 : Search time 136.222 Seconds

(without alignments)  
3438.621 Million cell updates/sec

Title: US-09-697-123b-22

Perfect score: 208

Sequence: 1 tcaagagaagcgtacgac.....ccgttgaaaccagcagcat 208

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapept 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_101002:\*

1: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:\*  
2: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:\*  
3: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:\*  
4: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:\*  
5: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:\*  
6: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:\*  
7: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:\*  
8: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:\*  
9: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:\*  
10: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:\*  
11: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:\*  
12: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:\*  
13: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:\*  
14: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:\*  
15: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:\*  
16: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:\*  
17: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:\*  
18: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:\*  
19: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:\*  
20: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:\*  
21: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:\*  
22: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:\*  
23: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:\*  
24: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	208	100.0	208	22	AA505222
2	206.4	99.2	208	22	AA505205
3	206.4	99.2	208	22	AA505206
4	206.4	99.2	208	22	AA505216
5	206.4	99.2	970	17	AA505216
6	206.4	99.2	3519	22	AAH51976
7	206.4	99.2	3534	22	AAH02079
8	206.4	99.2	3853	21	AAH74651
9	206.4	99.2	3853	21	AAH89994

10	171.2	82.3	208	22	AA505210	Mycobacterium ulce
11	169.6	81.5	208	22	AA505203	Mycobacterium gord
12	169.6	81.5	208	22	AA505211	Mycobacterium mari
13	168	80.8	208	22	AA505215	Mycobacterium avu
14	164.8	79.2	208	22	AA505201	Mycobacterium gord
15	164.8	79.2	208	22	AA505208	Mycobacterium kans
16	164.8	79.2	208	22	AA505219	Mycobacterium fort
17	164.8	79.2	208	22	AA505223	Mycobacterium haem
18	163.2	78.5	208	22	AA505202	Mycobacterium gord
19	158.4	76.2	208	22	AA505217	Mycobacterium cela
20	157.6	75.8	207	22	AA505204	Mycobacterium gord
21	152.8	73.5	207	22	AA505212	Mycobacterium szul
22	150.4	72.3	208	22	AA505224	Mycobacterium xeno
23	147.2	70.8	3447	14	AA051532	M. lepre prob gene
24	145.6	70.0	205	22	AA505220	Mycobacterium intr
25	137.6	66.2	214	22	AA505214	Mycobacterium malin
26	136.6	65.7	223	22	AA505209	Mycobacterium scro
27	128	61.5	214	22	AA505221	Mycobacterium abs
28	124.6	59.9	211	22	AA505218	Mycobacterium flav
29	123.2	59.2	214	22	AA505207	Mycobacterium chel
30	122.2	58.8	223	22	AA505213	Mycobacterium gast
31	88	42.3	3495	22	AAH65512	C glutamicum codin
32	88	42.3	349980	22	AAH68525	C glutamicum codin
33	64.4	31.0	27426	23	AA59541	Propionibacterium
34	39	18.8	1041	21	AA251711	Burkholderia cepac
35	38.8	18.7	1412	17	AA251711	Burkholderia cepac
36	38.2	18.4	1041	21	AA251715	Thermus aquaticus
37	38	18.3	1119	22	AA251715	Burkholderia cepac
38	37.4	18.0	1041	21	AA251704	Streptomyces ambol
39	37.4	18.0	1041	21	AA251710	Burkholderia vietn
40	37.4	18.0	1041	21	AA251712	Burkholderia cepac
41	37.4	18.0	1041	21	AA251713	Burkholderia cepac
42	37.4	18.0	1041	21	AA251714	Burkholderia cepac
43	37.4	18.0	1044	21	AA251698	Burkholderia cepac
44	37	17.8	6085	18	AA70153	S. longisporoflavus
45	36	17.3	1894	18	AA76542	S. longisporoflavus

## ALIGNMENTS

RESULT 1

AA505222

ID AA505222 standard; DNA: 208 BP.

AC AA505222;

DT 07-SEP-2001 (first entry)

XX

XX

DE Mycobacterium africanum rpoB gene fragment.

XX

KM Non-tuberculous mycobacteria: rpoB gene fragment; NTM; HIV; PRA: RFLP.

KW PCR-restriction fragment length polymorphism analysis; ds.

XX

OS Mycobacterium africanum.

XX

PN WO200131061-A1.

PD 03-MAY-2001.

XX

PF 27-OCT-2000; 2000WO-KR01223.

XX

PR 27-OCT-1999; 99KR-0046795.

XX

PA (ERUM-) ERUME BIOTECH CO LTD.

XX

PI Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;

XX

DR WPI: 2001-300520/31.

XX

PT New DNA fragments from the rpoB gene of mycobacteria, useful for

PT diagnosis and identification of many mycobacterial species by

PT restriction fragment length polymorphism -

XX Claim 1; Page 47; 50pp: English.  
XX  
XX The present sequence for *Mycobacterium africanum* rpoB gene  
CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from  
CC various *Mycobacterium* species. These rpoB gene fragments can be used  
CC in the diagnosis and identification of *Mycobacterium* species using a  
CC novel PCR-restriction fragment length polymorphism analysis (RFLP)  
CC method. The method comprises obtaining a restriction fragment length  
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,  
CC amplifying and digesting the DNA fragment from the microorganism to  
CC be identified and comparing the RFLP patterns from the known rpoB gene  
CC fragments with the unidentified fragment. The rpoB gene fragments  
CC are useful to identify a wide range of *Mycobacterium* species, e.g. for  
CC diagnosis or to obtain epidemiological and pathogenesis information for  
CC selection of appropriate therapies, including *M. tuberculosis*, *M. leprae*  
CC and non-tuberculous *Mycobacterium* (NTM) encountered in subjects infected  
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene  
CC fragments is rapid, precise, simple and cost effective (only 1 PCR  
CC required), and can differentiate between many species in a single  
CC experiment, including those difficult to distinguish by usual biochemical  
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for  
CC detecting specific *Mycobacterium* species.  
XX  
XX Sequence 208 BP; 46 A; 61 C; 67 G; 34 T; 0 other;  
SQ  
Query Match 100.0%; Score 208; DB 22; Length 208;  
Best Local Similarity 100.0%; Pred. No. 5, 2e-46;  
Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX  
XX 1 TCAGAGAGAACGCTACGACCTGCGCGCTGCTGCTATTAAGTCAACAGAAAGCTCG 60  
Db 1 TCAGAGAGAACGCTACGACCTGCGCGCTGCTGCTATTAAGTCAACAGAAAGCTCG 60  
XX  
XX 61 GCGTCATGTCGGCGAGCCCATCACTGCTGAGAGCTGACCGAAGAGAGCTCGCCCA 120  
Db 61 GCGTCATGTCGGCGAGCCCATCACTGCTGAGAGCTGACCGAAGAGAGCTCGCCCA 120  
XX  
XX 121 CCATGCAATATCTGTCGCTTGCACGAGGTCAGACCAAGATGATGCTTCGGGCGCGC 180  
Db 121 CCATGCAATATCTGTCGCTTGCACGAGGTCAGACCAAGATGATGCTTCGGGCGCGC 180  
XX  
XX 181 TCGAGGTGCGCGGTGGAACCGACGACAT 208  
Db 181 TCGAGGTGCGCGGTGGAACCGACGACAT 208  
XX  
XX  
XX RESULT 2  
XX AAS05205  
XX ID AAS05205 standard; DNA: 208 BP.  
XX AC  
XX AAS05205;  
XX  
XX 07-SEP-2001 (first entry)  
XX  
XX *Mycobacterium tuberculosis* rpoB gene fragment.  
XX  
XX Non-tuberculous *Mycobacterium*; rpoB gene fragment; NTM; HIV; PRA; RFLP;  
XX PCR-restriction fragment length polymorphism analysis; ds.  
XX  
XX *Mycobacterium tuberculosis*.  
XX  
XX WO200131061-A1.  
XX  
XX 03-MAY-2001.  
XX  
XX 27-OCT-2000; 2000WO-KR01223.  
XX  
XX 27-OCT-1999; 99KR-0046795.  
XX  
XX (ERUM-) ERUM BIOTECH CO LTD.  
XX  
XX Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;  
XX  
XX

XX  
XX MPI: 2001-300520/31.  
XX  
XX New DNA fragments from the rpoB gene of *Mycobacterium*, useful for  
PT diagnosis and identification of many *Mycobacterium* species by  
PT restriction fragment length polymorphism -  
XX  
XX Disclosure; Page 41; 50pp: English.  
XX  
XX The present sequence for *Mycobacterium tuberculosis* rpoB gene  
CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from  
CC various *Mycobacterium* species. These rpoB gene fragments can be used  
CC in the diagnosis and identification of *Mycobacterium* species using a  
CC novel PCR-restriction fragment length polymorphism analysis (RFLP)  
CC method. The method comprises obtaining a restriction fragment length  
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,  
CC amplifying and digesting the DNA fragment from the microorganism to  
CC be identified and comparing the RFLP patterns from the known rpoB gene  
CC fragments with the unidentified fragment. The rpoB gene fragments  
CC are useful to identify a wide range of *Mycobacterium* species, e.g. for  
CC diagnosis or to obtain epidemiological and pathogenesis information for  
CC selection of appropriate therapies, including *M. tuberculosis*, *M. leprae*  
CC and non-tuberculous *Mycobacterium* (NTM) encountered in subjects infected  
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene  
CC fragments is rapid, precise, simple and cost effective (only 1 PCR  
CC required), and can differentiate between many species in a single  
CC experiment, including those difficult to distinguish by usual biochemical  
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for  
CC detecting specific *Mycobacterium* species.  
XX  
XX Sequence 208 BP; 46 A; 62 C; 67 G; 33 T; 0 other;  
SQ  
Query Match 99.2%; Score 206.4; DB 22; Length 208;  
Best Local Similarity 99.5%; Pred. No. 1.6e-45;  
Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
XX  
XX 1 TCAGAGAGAACGCTACGACCTGCGCGCTGCTGCTATTAAGTCAACAGAAAGCTCG 60  
Db 1 TCAGAGAGAACGCTACGACCTGCGCGCTGCTGCTATTAAGTCAACAGAAAGCTCG 60  
XX  
XX 61 GCGTCATGTCGGCGAGCCCATCACTGCTGAGAGCTGACCGAAGAGAGCTCGCCCA 120  
Db 61 GCGTCATGTCGGCGAGCCCATCACTGCTGAGAGCTGACCGAAGAGAGCTCGCCCA 120  
XX  
XX 121 CCATGCAATATCTGTCGCTTGCACGAGGTCAGACCAAGATGATGCTTCGGGCGCGC 180  
Db 121 CCATGCAATATCTGTCGCTTGCACGAGGTCAGACCAAGATGATGCTTCGGGCGCGC 180  
XX  
XX 181 TCGAGGTGCGCGGTGGAACCGACGACAT 208  
Db 181 TCGAGGTGCGCGGTGGAACCGACGACAT 208  
XX  
XX  
XX RESULT 3  
XX AAS05206  
XX ID AAS05206 standard; DNA: 208 BP.  
XX AC  
XX AAS05206;  
XX  
XX 07-SEP-2001 (first entry)  
XX  
XX *Mycobacterium terrae* rpoB gene fragment.  
XX  
XX Non-tuberculous *Mycobacterium*; rpoB gene fragment; NTM; HIV; PRA; RFLP;  
XX PCR-restriction fragment length polymorphism analysis; ds.  
XX  
XX *Mycobacterium terrae*.  
XX  
XX WO200131061-A1.  
XX  
XX 03-MAY-2001.  
XX  
XX 27-OCT-2000; 2000WO-KR01223.  
XX  
XX

```
XX 27-OCT-1999; 99KR-0046795.
XX (ERUM-) ERUME BIOTECH CO LTD.
XX Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ:
XX WP1; 2001-300520/31.
XX
XX New DNA fragments from the rpoB gene of mycobacteria, useful for
XX diagnosis and identification of many mycobacterial species by
XX restriction fragment length polymorphism
XX
XX Claim 1; Page 42; 50pp; English.
XX
XX The present sequence for Mycobacterium terrae rpoB gene
XX fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
XX various Mycobacterium species. These rpoB gene fragments can be used
XX in the diagnosis and identification of Mycobacterium species using a
XX novel PCR-restriction fragment length polymorphism analysis (PRA)
XX method. The method comprises obtaining a restriction fragment length
XX polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
XX amplifying and digesting the DNA fragment from the microorganism to
XX be identified and comparing the RFLP patterns from the known rpoB gene
XX fragments with the unidentified fragment. The rpoB gene fragments
XX are useful to identify a wide range of Mycobacterium species, e.g. for
XX selection of appropriate epidemiological and pathogenesis information for
XX diagnosis or to obtain epidemiological and pathogenesis information for
XX selection of appropriate therapies, including M. tuberculosis, M. leprae
XX and non-tuberculous mycobacteria (NTM) encountered in subjects infected
XX with human immunodeficiency virus (HIV). Analysis of the rpoB gene
XX fragments is rapid, precise, simple and cost effective (only 1 PCR
XX required), and can differentiate between many species in a single
XX experiment, including those difficult to distinguish by usual biochemical
XX tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
XX detecting specific Mycobacterium species.
XX
XX Sequence 208 BP; 46 A; 62 C; 67 G; 33 T; 0 other:
XX
XX Query Match 99.2%; Score 206.4; DB 22; Length 208;
XX Best Local Similarity 99.5%; Pred. No. 1.6e-45;
XX Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX 1 TCAAGGAGAACGCGCTACGACCTGCGCGGCTGCGTATTAAGTCAACAAGAGCTCG 60
XX |
XX 1 TCAAGGAGAACGCGCTACGACCTGCGCGGCTGCGTATTAAGTCAACAAGAGCTCG 60
XX
XX 61 GGCTGCATGTGCGGCGAGCCCATCAGCTGCTGACGCTGACCGAAGAGAGCTGTGGCCA 120
XX |
XX 61 GGCTGCATGTGCGGCGAGCCCATCAGCTGCTGACGCTGACCGAAGAGAGCTGTGGCCA 120
XX
XX 121 CCATCGAATATCTGTCGCTTGCACGAGGTCAGACGACGATGATCGTTCCGGCGCGC 180
XX |
XX 121 CCATCGAATATCTGTCGCTTGCACGAGGTCAGACGACGATGATCGTTCCGGCGCGC 180
XX
XX 181 TCGAGGTGCGCGGTGGAACCGAGCAGCAT 208
XX |
XX 181 TCGAGGTGCGCGGTGGAACCGAGCAGCAT 208
XX
XX
XX RESULT 4
XX AAS05216
XX ID AAS05216 standard; DNA; 208 BP.
XX AC AAS05216;
XX
XX 07-SEP-2001 (first entry)
XX
XX Mycobacterium bovis rpoB gene fragment.
XX
XX Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;
XX PCR-restriction fragment length polymorphism analysis; ds.
XX
XX Mycobacterium bovis.
```

```
XX WO200131061-A1.
XX
XX 03-MAY-2001.
XX
XX 27-OCT-2000; 2000MO-KR01223.
XX
XX 27-OCT-1999; 99KR-0046795.
XX
XX (ERUM-) ERUME BIOTECH CO LTD.
XX Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ:
XX WP1; 2001-300520/31.
XX
XX New DNA fragments from the rpoB gene of mycobacteria, useful for
XX diagnosis and identification of many mycobacterial species by
XX restriction fragment length polymorphism
XX
XX Claim 1; Page 45; 50pp; English.
XX
XX The present sequence for Mycobacterium bovis rpoB gene
XX fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
XX various Mycobacterium species. These rpoB gene fragments can be used
XX in the diagnosis and identification of Mycobacterium species using a
XX novel PCR-restriction fragment length polymorphism analysis (PRA)
XX method. The method comprises obtaining a restriction fragment length
XX polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
XX amplifying and digesting the DNA fragment from the microorganism to
XX be identified and comparing the RFLP patterns from the known rpoB gene
XX fragments with the unidentified fragment. The rpoB gene fragments
XX are useful to identify a wide range of Mycobacterium species, e.g. for
XX diagnosis or to obtain epidemiological and pathogenesis information for
XX selection of appropriate therapies, including M. tuberculosis, M. leprae
XX and non-tuberculous mycobacteria (NTM) encountered in subjects infected
XX with human immunodeficiency virus (HIV). Analysis of the rpoB gene
XX fragments is rapid, precise, simple and cost effective (only 1 PCR
XX required), and can differentiate between many species in a single
XX experiment, including those difficult to distinguish by usual biochemical
XX tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
XX detecting specific Mycobacterium species.
XX
XX Sequence 208 BP; 46 A; 62 C; 67 G; 33 T; 0 other:
XX
XX Query Match 99.2%; Score 206.4; DB 22; Length 208;
XX Best Local Similarity 99.5%; Pred. No. 1.6e-45;
XX Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX 1 TCAAGGAGAACGCGCTACGACCTGCGCGGCTGCGTATTAAGTCAACAAGAGCTCG 60
XX |
XX 1 TCAAGGAGAACGCGCTACGACCTGCGCGGCTGCGTATTAAGTCAACAAGAGCTCG 60
XX
XX 61 GGCTGCATGTGCGGCGAGCCCATCAGCTGCTGACGCTGACCGAAGAGAGCTGTGGCCA 120
XX |
XX 61 GGCTGCATGTGCGGCGAGCCCATCAGCTGCTGACGCTGACCGAAGAGAGCTGTGGCCA 120
XX
XX 121 CCATCGAATATCTGTCGCTTGCACGAGGTCAGACGACGATGATCGTTCCGGCGCGC 180
XX |
XX 121 CCATCGAATATCTGTCGCTTGCACGAGGTCAGACGACGATGATCGTTCCGGCGCGC 180
XX
XX 181 TCGAGGTGCGCGGTGGAACCGAGCAGCAT 208
XX |
XX 181 TCGAGGTGCGCGGTGGAACCGAGCAGCAT 208
XX
XX
XX RESULT 5
XX AAT09676
XX ID AAT09676 standard; DNA; 970 BP.
XX AC AAT09676;
XX
XX 15-OCT-1996 (first entry)
```

```
DE  Mycobacterium tuberculosis rpoB gene DNA sequence.
XX
KW  Tuberculosis: disease diagnosis: Oligonucleotide: DNA primer: PCR;
KW  polymerase chain reaction: DNA amplification: rpoB locus: TB; ss.
XX
OS  Mycobacterium tuberculosis.
XX
Key  Location/Qualifiers
FH  10..27
FH  primer_bind
FT  /tag= a
FT  /note= "primer FENLEF"
FT  226..243
FT  primer_bind
FT  /tag= b
FT  /note= "primer DDIDL"
FT  226..240
FT  primer_bind
FT  /tag= c
FT  /note= "primer DDIDH"
FT  338..364
FT  primer_bind
FT  /tag= d
FT  /note= "primer rpo95"
FT  348..373
FT  primer_bind
FT  /tag= e
FT  /note= "primer rpo105"
FT  354..373
FT  primer_bind
FT  /tag= f
FT  /note= "primer KY290"
FT  372..373
FT  misc_feature
FT  /tag= g
FT  /note= "M. tuberculosis signature nucleotide"
FT  433..434
FT  misc_feature
FT  /tag= h
FT  /note= "M. tuberculosis signature nucleotide"
FT  438
FT  misc_feature
FT  /tag= i
FT  /note= "M. tuberculosis signature nucleotide"
FT  468..469
FT  misc_feature
FT  /tag= j
FT  /note= "M. tuberculosis signature nucleotide"
FT  486
FT  misc_feature
FT  /tag= k
FT  /note= "M. tuberculosis signature nucleotide"
FT  501
FT  misc_feature
FT  /tag= l
FT  /note= "M. tuberculosis signature nucleotide"
FT  516
FT  misc_feature
FT  /tag= m
FT  /note= "M. tuberculosis signature nucleotide"
FT  516..535
FT  primer_bind
FT  /tag= n
FT  /note= "primer rpo273"
FT  525
FT  misc_feature
FT  /tag= o
FT  /note= "M. tuberculosis signature nucleotide"
FT  525..541
FT  primer_bind
FT  /tag= p
FT  /note= "primer KY292"
FT  536..562
FT  primer_bind
FT  /tag= q
FT  /note= "primer rpo293"
FT  640..666
FT  primer_bind
FT  /tag= r
FT  /note= "primer rpo397"
FT  952..966
FT  primer_bind
FT  /tag= s
FT  /note= "primer NMQR0-1"
FT  952..966
FT  primer_bind
FT  /tag= t
FT  /note= "primer NMQR0-2"
XX
XX  WO9533074-A1.
XX
XX  07-DEC-1995.
XX
```

```
PF  26-MAY-1995: 95WO-US06790.
XX
XX  26-MAY-1994: 94US-0250030.
XX
XX  (HOFF ) HOFFMANN LA ROCHE INC.
XX  (MAYO-) MAYO FOUNDATION.
XX
XX  Felmlae TA, Hunt JM, Persing DH, Roberts GD, Whelen AC;
XX  Young KKY;
XX
XX  WPL: 1996-030581/03.
XX
XX  Detection of Mycobacterium tuberculosis - by amplifying sample DNA
XX  with a primer set that targets portions of the gene encoding rpoB.
XX
XX  Disclosure: Fig.3: 54pp: English.
```

```
CC  This oligonucleotide DNA primer is specific for Mycobacterium
CC  tuberculosis, and may be used to amplify a sample DNA by targeting
CC  a portion of the gene encoding rpoB. The 1st several bases comprise a
CC  nonhybridizing tail consisting of filler bases followed by
CC  a restriction site incorporated to facilitate cloning using the
CC  amplicon at a later date, if desired. The remaining bases hybridize
CC  to bacterial rpoB DNA. The method provides for the detection of M.
CC  tuberculosis and the concurrent determination of its drug
CC  susceptibility, particularly to rifamycin. The method can provide
CC  often greater than 95% sensitivity and 100% specificity. The
CC  biological sample is a fluid or tissue sample from a human.
XX
```

Sequence 970 BP: 182 A: 302 C: 330 G: 156 T: 0 other:

Query Match 99.2%: Score 206.4: DB 17: Length 970:  
Best Local Similarity 99.5%: Pred. No. 2e-45;  
Matches 207: Conservative 0: Mismatches 1: Indels 0: Gaps 0;

```
QY  1 TCAAGAGAAGCGCTACGACCTGCGCCGCTGCGCTATTAAGTCAACAAGAGCTCG 60
DB  |||
QY  26 TCAAGAGAAGAGCGCTACGACCTGCGCCGCTGCGCTATTAAGTCAACAAGAGCTCG 85
DB  |||
QY  61 GGCTGATGTCGGCGAGCCCATCAGTCGTCGACGCTGACCGAAGAGAGCTGCGCA 120
DB  |||
QY  86 GGCTGATGTCGGCGAGCCCATCAGTCGTCGACGCTGACCGAAGAGAGCTGCGCA 145
DB  |||
QY  121 CCATGATATGTCGGCTGCGACGAGGTCACAGACCATGATGTCGCGCGCG 180
DB  |||
QY  146 CCATGATATGTCGGCTGCGACGAGGTCACAGACCATGATGTCGCGCGCG 205
DB  |||
QY  181 TCGAGGTGCGGTGGAACCGAGACAT 208
DB  |||
QY  206 TCGAGGTGCGGTGGAACCGAGACAT 233
```

```
RESULT 6
AAH51976
ID  AAH51976 standard; DNA: 3519 BP.
XX
XX  AAH51976;
XX
XX  04-SEP-2001 (first entry)
XX
XX  Mycobacterium tuberculosis potential drug target gene SEQ ID 30.
XX
XX  Drug target; growth; organism viability; characterisation; ds.
XX
XX  Mycobacterium tuberculosis.
XX
XX  WO200135317-A1.
XX
XX  17-MAY-2001.
XX
XX  13-NOV-2000; 2000WO-US31152.
XX
XX  12-NOV-1999; 99US-0165086.
XX
```

PR 12-NOV-1999; 9905-0165124.  
PR 01-FEB-2000; 200005-0179531.  
XX  
XX  
XX (REGC ) UNIV CALIFORNIA.  
PI Eisenberg D, Rotstein SH, Marcotte EM.  
XX  
XX WPI: 2001-329193/34.  
DR P-PSDB: AAG81125.  
XX  
XX  
PT Identifying nucleotide or polypeptide sequence for use as drug target,  
PT between providing algorithm that analyzes a functional relationship  
PT between nucleotide or polypeptide sequences, and comparing the  
PT sequences -  
PS  
PS Disclosure: Page 68-69; 207pp: English.  
XX  
XX This invention relates to a method for identifying a nucleotide or  
XX polypeptide sequence that may be a drug target, or essential for growth  
XX or viability of an organism. Polynucleotide sequences AAH51947 - AAH52092  
XX represent DNA encoding proteins AAG81096 - AAG81241, Mycobacterium  
XX tuberculosis proteins which are potential drug targets. The DNA and  
XX protein sequences are used to illustrate the method of the invention. The  
XX method involves providing an unknown nucleotide or polypeptide sequence,  
XX and comparing it to a number of sequences along with at least one  
XX algorithm capable of analysing a functional relationship between  
XX nucleotide and polypeptide sequences. The method is useful for  
XX characterizing the function of nucleic acids and polypeptides that may be  
XX useful as a target for a drug or essential for the growth or viability of  
XX an organism.  
XX  
SQ Sequence 3519 BP; 675 A; 1081 C; 1183 G; 580 T; 0 other;  
Query Match 99.2%; Score 206.4; DB 22; Length 3519;  
Best Local Similarity 99.5%; Pred. No. 2.3e-45;  
Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 TCAAGAGAGAGCCCTAGACAGCTGGCCGGTGGCTGATATAGGTCACAGAGAGCTCG 60  
DB 884 TCAAGAGAGAGCCCTAGACAGCTGGCCGGTGGCTGATATAGGTCACAGAGAGCTCG 943  
QY 61 GCGTCGATGCGCGGAGAGCCCATCAGCTGTCGACGCTGACCGAAGACGTCGTGGCCA 120  
DB 944 GCGTCGATGCGCGGAGAGCCCATCAGCTGTCGACGCTGACCGAAGACGTCGTGGCCA 1003  
QY 121 CCATCGAATATCTGTCGCGGTGACAGAGGTCAGACCATGATCGTCCGGGGCGG 180  
DB 1004 CCATCGAATATCTGTCGCGGTGACAGAGGTCAGACCATGATCGTCCGGGGCGG 1063  
QY 181 TCGAGGTGCCGGTGAACCGACGACAT 208  
DB 1064 TCGAGGTGCCGGTGAACCGACGACAT 1091  
RESULT 7  
AAH02079  
ID AAH02079 standard; DNA; 3534 BP.  
XX  
XX AAH02079;  
AC  
XX  
XX 24-JUL-2001 (first entry)  
XX  
XX Mycobacterium tuberculosis nucleotide sequence SEQ ID NO:2072.  
XX  
XX Species specific; genus specific; family specific; probe; detection;  
XX identification; algal; archaeal; bacterial; fungal; parasitical;  
XX microorganism; diagnosis; translation elongation factor Tu; toxin;  
XX translation elongation factor G; RecA recombinase; resistance;  
XX catalytic subunit of proton-translocating ATPase; antimicrobial;  
XX vaccine; primer; ds.  
XX  
XX Mycobacterium tuberculosis.

PN W0200123604-A2.  
XX  
XX  
XX 05-APR-2001.  
XX  
XX 28-SEP-2000; 2000WO-CA01150.  
XX  
XX 28-SEP-1999; 99CA-2283458.  
PR 19-MAY-2000; 2000CA-2307010.  
XX  
XX  
XX (INFE-) INFECTIO DIAGNOSTIC (IDI) INC.  
XX  
XX Bergeron MG, Boissinot M, Huletsky A, Menard C, Ouellette M;  
XX Picard FJ, Roy PH;  
XX WPI: 2001-245006/25.  
XX  
XX  
XX Nucleic acid sequences are used to generate universal probes and  
XX primers which can be used to identify and detect the presence of algal,  
XX archaeal, bacterial, fungal and parasitical species in a test sample -  
XX  
XX Disclosure: Page 1478-1479; 1580pp: English.  
XX  
XX The present invention describes a method for generating a repertoire of  
XX nucleic acids of tuf, fus, atp and/or recA genes from which probes  
XX and/or primers are derived. The method comprises amplifying the nucleic  
XX acids of determined algal, archaeal, bacterial, fungal and parasitical  
XX species with a combination of defined primer pairs. The method can be  
XX used for producing probes and/or primers for detecting one or more  
XX related microorganisms e.g. algae, archaea, bacteria, fungi and  
XX parasites, for universal detection and for specific and ubiquitous  
XX detection and identification of an algal, archaeal, bacterial, fungal  
XX and parasitical species, genus, family and group. A nucleic acid (1)  
XX obtained using the method of the invention can be used for the universal  
XX detection of any bacterium, fungus or parasite in a sample and for the  
XX detection of at least one antimicrobial agent resistance gene or at  
XX least one toxin gene. hexa nucleic acids are used for the specific and  
XX ubiquitous detection and for identification of Streptococcus pneumoniae,  
XX microorganisms. Microbial species or genus or family or phylum or group  
XX which can be detected include Abiotrophia adiacens, Bordetella sp.,  
XX Corynebacterium sp., Enterobacteriaceae group, Escherichia coli,  
XX Mycobacteriaceae family, Pseudomonas group, Streptococcus sp.,  
XX Neisseria gonorrhoeae and Staphylococcus sp.. Using DNA based tests  
XX provides faster results than substrate specificity tests as results can  
XX be determined in an hour and improved accuracy is also achieved.  
XX AAH00010 to AAH002304 represent nucleotide sequences and primers/probes  
XX which are given in the exemplification of the present invention.  
XX  
SQ Sequence 3534 BP; 679 A; 1081 C; 1188 G; 586 T; 0 other;  
Query Match 99.2%; Score 206.4; DB 22; Length 3534;  
Best Local Similarity 99.5%; Pred. No. 2.3e-45;  
Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 TCAAGAGAGAGCCCTAGACAGCTGGCCGGTGGCTGATATAGGTCACAGAGAGCTCG 60  
DB 902 TCAAGAGAGAGCCCTAGACAGCTGGCCGGTGGCTGATATAGGTCACAGAGAGCTCG 961  
QY 61 GCGTCGATGCGCGGAGAGCCCATCAGCTGTCGACGCTGACCGAAGACGTCGTGGCCA 120  
DB 962 GCGTCGATGCGCGGAGAGCCCATCAGCTGTCGACGCTGACCGAAGACGTCGTGGCCA 1021  
QY 121 CCATCGAATATCTGTCGCGGTGACAGAGGTCAGACCATGATCGTCCGGGGCGG 180  
DB 1022 CCATCGAATATCTGTCGCGGTGACAGAGGTCAGACCATGATCGTCCGGGGCGG 1081  
QY 181 TCGAGGTGCCGGTGAACCGACGACAT 208  
DB 1082 TCGAGGTGCCGGTGAACCGACGACAT 1109  
RESULT 8  
AAA74651

ID AAA74651 standard; DNA; 3853 BP.  
XX  
AC AAA74651;  
XX  
DT 06-DEC-2000 (first entry)  
XX  
DE Mycobacterium tuberculosis rpoB gene.  
XX  
DE Mycobacterium tuberculosis rpoB gene.  
XX  
KM Mycobacterium tuberculosis; rpoB; RNA polymerase beta subunit;  
XX rifampin resistance; mutation detection; ds.  
XX  
OS Mycobacterium tuberculosis.  
XX  
PN WO200043546-A2.  
XX  
PD 27-JUL-2000.  
XX  
PE 20-DEC-1999; 99WO-US30377.  
XX  
PR 19-JAN-1999; 99US-0233996.  
XX  
PA (DADE-) DADE BEHRING INC.  
XX  
PI Liu YP, Kurn N;  
XX  
DR WPI: 2000-524243/47.  
XX  
PT Method for detecting drug resistance in a strain of an organism,  
XX particularly for detecting rifampin resistance in Mycobacterium  
XX tuberculosis -  
XX  
PS Example 1; Fig 4; 86pp; English.  
XX  
CC The present sequence is the rpoB gene from Mycobacterium tuberculosis.  
CC Rifampin resistance is largely associated with point mutations  
CC localised in a small core region of 81 base pairs in the rpoB gene, which  
CC encodes the RNA polymerase beta subunit. To detect a mutation, a complex  
CC is formed comprising a first sequence representing the predetermined  
CC region of the gene of the organism and a second sequence representing the  
CC corresponding region of the gene of the wild type organism in double  
CC stranded form. Each member of at least one pair of non-complementary  
CC strands within the complex has a label. The association of the labels in  
CC the complex is related to the presence of the mutation. The presence of  
CC the mutation is related to the drug resistance of the strain.  
XX  
SQ Sequence 3853 BP; 723 A; 1173 C; 1293 G; 664 T; 0 other;  
Query Match 99.2%; Score 206.4; DB 21; Length 3853;  
Best Local Similarity 99.5%; Pred. No. 2.3e-45;  
Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 TCAAGGAGAAAGCGCTACGACCTGGCCCGCGTGGCTATTAAGGTCAACAAGAAGCTCG 60  
DB 1477 TCAAGGAGAAAGCGCTACGACCTGGCCCGCGTGGCTATTAAGGTCAACAAGAAGCTCG 1536  
QY 61 GGCTGATGTCGGCGAGCCATCAGCTGCTGACCGTGAACGAGAAAGACGTCGGCCA 120  
DB 1537 GGCTGATGTCGGCGAGCCATCAGCTGCTGACCGTGAACGAGAAAGACGTCGGCCA 1596  
QY 121 CCATCGAATATCTGTCGCTTGCACAGAGGTCAAGACCAAGATGATCTTCCGGCGCGC 180  
DB 1597 CCATCGAATATCTGTCGCTTGCACAGAGGTCAAGACCAAGATGATCTTCCGGCGCGC 180  
QY 181 TCGAGGTGCGCGGTGAACCGACGACAT 208  
DB 1657 TCGAGGTGCGCGGTGAACCGACGACAT 1684

RESULT 9  
ID AAA89994 standard; DNA; 3853 BP.  
XX  
AC AAA89994;  
XX

XX  
DT 18-DEC-2000 (first entry)  
XX  
DE M. tuberculosis rpoB gene encoding RNA polymerase beta subunit.  
XX  
DE M. tuberculosis rpoB gene encoding RNA polymerase beta subunit.  
XX  
KM Drug resistance; rifampin; rifampicin; pyrazinamide; rpoB;  
XX RNA polymerase beta subunit; ds.  
XX  
OS Mycobacterium tuberculosis.  
XX  
PN WO200043545-A2.  
XX  
PD 27-JUL-2000.  
XX  
PE 14-DEC-1999; 99WO-US29517.  
XX  
PR 19-JAN-1999; 99US-0233996.  
XX  
PR 22-APR-1999; 99US-0296894.  
XX  
PA (DADE-) DADE BEHRING INC.  
XX  
PI Liu YP, Kurn N;  
XX  
DR WPI: 2000-499235/44.  
XX  
PT Detecting resistance of drugs such as rifampicin in strains of  
XX Mycobacterium, comprising detecting mutations in a gene and relating  
XX them to drug resistance -  
XX  
PS Example 1; Fig 4; 91pp; English.  
XX  
CC This invention relates to a method for detecting drug resistance in a  
CC strain of an organism. The method comprises detecting the presence of at  
CC least 1 mutation in a first sequence and relating the presence of the  
CC mutation to drug resistance. Included in the invention are a kit for  
CC carrying out the method and a method for detecting the presence of a  
CC difference between two related nucleic acid sequences in an organism. The  
CC methods are useful for detecting resistance to drugs such as rifampin and  
CC pyrazinamide in Mycobacterium. The present sequence represents the  
CC Mycobacterium tuberculosis rpoB gene (which encodes the RNA polymerase  
CC beta subunit). The sequence is used in an example of the method of the  
CC invention for the detection of rifampin resistance in M. tuberculosis.  
XX  
SQ Sequence 3853 BP; 723 A; 1173 C; 1293 G; 664 T; 0 other;  
Query Match 99.2%; Score 206.4; DB 21; Length 3853;  
Best Local Similarity 99.5%; Pred. No. 2.3e-45;  
Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 TCAAGGAGAAAGCGCTACGACCTGGCCCGCGTGGCTATTAAGGTCAACAAGAAGCTCG 60  
DB 1477 TCAAGGAGAAAGCGCTACGACCTGGCCCGCGTGGCTATTAAGGTCAACAAGAAGCTCG 1536  
QY 61 GGCTGATGTCGGCGAGCCATCAGCTGCTGACCGTGAACGAGAAAGACGTCGGCCA 120  
DB 1537 GGCTGATGTCGGCGAGCCATCAGCTGCTGACCGTGAACGAGAAAGACGTCGGCCA 1596  
QY 121 CCATCGAATATCTGTCGCTTGCACAGAGGTCAAGACCAAGATGATCTTCCGGCGCGC 180  
DB 1597 CCATCGAATATCTGTCGCTTGCACAGAGGTCAAGACCAAGATGATCTTCCGGCGCGC 180  
QY 181 TCGAGGTGCGCGGTGAACCGACGACAT 208  
DB 1657 TCGAGGTGCGCGGTGAACCGACGACAT 1684

RESULT 10  
ID AAS05210 standard; DNA; 208 BP.  
XX  
AC AAS05210;  
XX  
DT 07-SEP-2001 (first entry)



XX DE Mycobacterium ulcerans rpoB gene fragment.  
XX KM Non-tuberculous mycobacteria: rpoB gene fragment; NTM; HIV; PRA; RFLP;  
XX KM PCR-restriction fragment length polymorphism analysis; ds.  
XX OS Mycobacterium ulcerans.  
XX PN WO200131061-A1.  
XX PD 03-MAY-2001.  
XX PF 27-OCT-2000: 2000MO-KR01223.  
XX PR 27-OCT-1999: 99KR-0046795.  
XX PA (ERUM-) ERUME BIOTECH CO LTD.  
XX PI Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;  
XX DR WPI: 2001-300520/31.  
XX PT New DNA fragments from the rpoB gene of mycobacteria, useful for  
XX PT diagnosis and identification of many mycobacterial species by  
XX PT restriction fragment length polymorphism -  
XX PS Claim 1: Page 43: 50pp; English.  
XX CC The present sequence for Mycobacterium ulcerans rpoB gene  
XX CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from  
XX CC various Mycobacterial species. These rpoB gene fragments can be used  
XX CC in the diagnosis and identification of Mycobacterium species using a  
XX CC novel PCR-restriction fragment length polymorphism analysis (PRA)  
XX CC method. The method comprises obtaining a restriction fragment length  
XX CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,  
XX CC amplifying and digesting the DNA fragment from the microorganism to  
XX CC be identified and comparing the RFLP patterns from the known rpoB gene  
XX CC fragments with the unidentified fragment. The rpoB gene fragments  
XX CC are useful to identify a wide range of Mycobacterium species, e.g. for  
XX CC diagnosis or to obtain epidemiological and pathogenesis information for  
XX CC selection of appropriate therapies, including M. tuberculosis, M. lepre  
XX CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected  
XX CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene  
XX CC fragments is rapid, precise, simple and cost effective (only 1 PCR  
XX CC required), and can differentiate between many species in a single  
XX CC experiment, including those difficult to distinguish by usual biochemical  
XX CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for  
XX CC detecting specific Mycobacterial species.  
XX SO Sequence 208 BP; 45 A; 68 C; 67 G; 28 T; 0 other;  
Query Match 82.3%; Score 171.2; DB 22: Length 208;  
Best Local Similarity 88.9%; Pred. No. 3.4e-36;  
Matches 185; Conservative 0; Mismatches 23; Indels 0; Gaps 0;  
OY 1 TCAAGGAGAGCGGTACGACCTGGCCCGCTGCTGCTTAAGTCAACAGAGCTCG 60  
DB 1 TCAAGGAGAGCGGTACGACCTGGCTGCTGCTGCTGCTCAAGGTCACAAAGAGCTCG 60  
OY 61 GCGTCATGTCGGCGAGCCCATCACTGCTGAGCGTGAAGGAGAGAGCTGGCGCA 120  
DB 61 GCGTCGAAAGCGCGGCGAGCCCATCACTGAGCTGAGCGGAGAGAGAGCTGGCGCA 120  
OY 121 CCATCGAATATCTGCTCGCTTGCAGAGGGTCAAGCAGATGATCTTCCGGCGGCG 180  
DB 121 CCATCGAATATCTGCTCGCTTGCAGAGGGTCAAGCAGATGATGACCGGCTCCGGGGG 180  
OY 181 TCGAGTGGCGGTGAGAAACGAGACAT 208  
DB 181 TCGAGTGGCGGTGAGAAACGAGACAT 208  
RESULT 11

AAS05203  
ID AAS05203 standard; DNA; 208 BP.  
XX AC AAS05203;  
XX AC AAS05203;  
XX DT 07-SEP-2001 (first entry)  
XX KM Mycobacterium gordonae type III rpoB gene fragment.  
XX DE Mycobacterium gordonae type III rpoB gene fragment.  
XX KM Non-tuberculous mycobacteria: rpoB gene fragment; NTM; HIV; PRA; RFLP;  
XX KM PCR-restriction fragment length polymorphism analysis; ds.  
XX OS Mycobacterium gordonae type III.  
XX PN WO200131061-A1.  
XX PD 03-MAY-2001.  
XX PF 27-OCT-2000: 2000MO-KR01223.  
XX PR 27-OCT-1999: 99KR-0046795.  
XX PA (ERUM-) ERUME BIOTECH CO LTD.  
XX PI Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;  
XX DR WPI: 2001-300520/31.  
XX PT New DNA fragments from the rpoB gene of mycobacteria, useful for  
XX PT diagnosis and identification of many mycobacterial species by  
XX PT restriction fragment length polymorphism -  
XX PS Claim 1: Page 41: 50pp; English.  
XX CC The present sequence for Mycobacterium gordonae type III rpoB gene  
XX CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from  
XX CC various Mycobacterial species. These rpoB gene fragments can be used  
XX CC in the diagnosis and identification of Mycobacterium species using a  
XX CC novel PCR-restriction fragment length polymorphism analysis (PRA)  
XX CC method. The method comprises obtaining a restriction fragment length  
XX CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,  
XX CC amplifying and digesting the DNA fragment from the microorganism to  
XX CC be identified and comparing the RFLP patterns from the known rpoB gene  
XX CC fragments with the unidentified fragment. The rpoB gene fragments  
XX CC are useful to identify a wide range of Mycobacterium species, e.g. for  
XX CC diagnosis or to obtain epidemiological and pathogenesis information for  
XX CC selection of appropriate therapies, including M. tuberculosis, M. lepre  
XX CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected  
XX CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene  
XX CC fragments is rapid, precise, simple and cost effective (only 1 PCR  
XX CC required), and can differentiate between many species in a single  
XX CC experiment, including those difficult to distinguish by usual biochemical  
XX CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for  
XX CC detecting specific Mycobacterial species.  
XX SO Sequence 208 BP; 46 A; 69 C; 63 G; 30 T; 0 other;  
Query Match 81.5%; Score 169.6; DB 22: Length 208;  
Best Local Similarity 88.5%; Pred. No. 8.9e-36;  
Matches 184; Conservative 0; Mismatches 24; Indels 0; Gaps 0;  
OY 1 TCAAGGAGAGCGGTACGACCTGGCCCGCTGCTGCTTAAGTCAACAGAGCTCG 60  
DB 1 TCAAGGAGAGCGGTACGACCTGGCTGCTGCTGCTGCTCAAGGTCACAAAGAGCTCG 60  
OY 61 GCGTCATGTCGGCGAGCCCATCACTGCTGAGCGTGAAGGAGAGAGCTGGCGCA 120  
DB 61 GCGTCGACGTCGGGAGATCGATCACCAGCTGACCGGAGAGAGAGCTGGCGCA 120  
OY 121 CCATCGAATATCTGCTCGCTTGCAGAGGGTCAAGCAGATGATCTTCCGGCGGCG 180  
DB 121 CCATCGAATATCTGCTCGCTTGCAGAGGGTCAAGCAGATGATGACCGGCTCCGGGGG 180

OY 181 TCGAGTGCCTGGTGAACCGACGACAT 208  
||||| ||||||| ||||||| |||||||  
Db 181 CCGAGGTCCTCCGTGACACGACGACAT 208

## RESULT 12

AAS05211  
ID AAS05211 standard; DNA: 208 BP.

AC AAS05211;

DT 07-SEP-2001 (first entry)

DE Mycobacterium marinum rpoB gene fragment.

KM Non-tuberculous mycobacteria: rpoB gene fragment; NTM; HIV; PRA; RFLP;  
KM PCR-restriction fragment length polymorphism analysis; ds.

OS Mycobacterium marinum.

PN WO200131061-A1.

PD 03-MAY-2001.

PF 27-OCT-2000; 2000WO-KR01223.

PR 27-OCT-1999; 99KR-0046795.

PA (ERUM-) ERUME BIOTECH CO LTD.

PI Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;  
XN WPI: 2001-300520/31.

PT New DNA fragments from the rpoB gene of mycobacteria, useful for  
PT diagnosis and identification of many mycobacterial species by  
PT restriction fragment length polymorphism -

PS Claim 1; Page 43; 50pp; English.

CC The present sequence for Mycobacterium marinum rpoB gene  
CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from  
CC various Mycobacterial species. These rpoB gene fragments can be used  
CC in the diagnosis and identification of Mycobacterium species using a  
CC novel PCR-restriction fragment length polymorphism analysis (PRA)  
CC method. The method comprises obtaining a restriction fragment length  
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,  
CC amplifying and digesting the DNA fragment from the microorganism to  
CC be identified and comparing the RFLP patterns from the known rpoB gene  
CC fragments with the unidentified fragment. The rpoB gene fragments  
CC are useful to identify a wide range of Mycobacterium species, e.g. for  
CC diagnosis or to obtain epidemiological and pathogenesis information for  
CC selection of appropriate therapies, including M. tuberculosis, M. leprae  
CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected  
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene  
CC fragments is rapid, precise, simple and cost effective (only 1 PCR  
CC experiment), and can differentiate between many species in a single  
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for  
CC detecting specific Mycobacterial species.

XX Sequence 208 BP; 45 A; 69 C; 68 G; 26 T; 0 other;

Query Match 81.5%; Score 169.6; DB 22; Length 208;

Best Local Similarity 88.5%; Pred. No. 8.9e-36;

Matches 184; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

OY 1 TCAAGAGAGCGCTACGACCTGCGCGCTGCTGCTAAGGTCACACAGAGCTCG 60  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 1 TCAAGAGAGAGCGCTACGACCTGCGCGCGCTGCTGCTAAGGTCACACAGAGCTCG 60

OY 61 GCGTCACGTCGGGCGAGCCCATCTACGTCGACCTGACCCGAAGAAGAGCTGTCGCA 120  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 61 GCGTCACGTCGGGCGAGCCCATCTACGTCGACCTGACCCGAAGAAGAGCTGTCGCA 120  
OY 121 CCATCGAATATCTGACCCGCTGACAGGCTCAGACACGATGATCGTCCGGCGCGC 180  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
Db 121 CCATCGAATATCTGTCGCTTGCACGAGGCGCCGACCGCATGACCGCTCCGGCGCGC 180  
OY 181 TCGAGTGCCTGGTGAACCGACGACAT 208  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
Db 181 TCGAGTGCCTGGTGAACCGACGACGACAT 208

## RESULT 13

AAS05215  
ID AAS05215 standard; DNA: 208 BP.

AC AAS05215;

DT 07-SEP-2001 (first entry)

DE Mycobacterium avium rpoB gene fragment.

KM Non-tuberculous mycobacteria: rpoB gene fragment; NTM; HIV; PRA; RFLP;  
KM PCR-restriction fragment length polymorphism analysis; ds.

OS Mycobacterium avium.

PN WO200131061-A1.

PD 03-MAY-2001.

PF 27-OCT-2000; 2000WO-KR01223.

PR 27-OCT-1999; 99KR-0046795.

PA (ERUM-) ERUME BIOTECH CO LTD.

PI Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;  
XN WPI: 2001-300520/31.

PT New DNA fragments from the rpoB gene of mycobacteria, useful for  
PT diagnosis and identification of many mycobacterial species by  
PT restriction fragment length polymorphism -

PS Claim 1; Page 44; 50pp; English.

CC The present sequence for Mycobacterium avium rpoB gene  
CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from  
CC various Mycobacterial species. These rpoB gene fragments can be used  
CC in the diagnosis and identification of Mycobacterium species using a  
CC novel PCR-restriction fragment length polymorphism analysis (PRA)  
CC method. The method comprises obtaining a restriction fragment length  
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,  
CC amplifying and digesting the DNA fragment from the microorganism to  
CC be identified and comparing the RFLP patterns from the known rpoB gene  
CC fragments with the unidentified fragment. The rpoB gene fragments  
CC are useful to identify a wide range of Mycobacterium species, e.g. for  
CC diagnosis or to obtain epidemiological and pathogenesis information for  
CC selection of appropriate therapies, including M. tuberculosis, M. leprae  
CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected  
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene  
CC fragments is rapid, precise, simple and cost effective (only 1 PCR  
CC experiment), and can differentiate between many species in a single  
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for  
CC detecting specific Mycobacterial species.

XX Sequence 208 BP; 44 A; 69 C; 69 G; 26 T; 0 other;

Query Match 80.8%; Score 168; DB 22; Length 208;

Best Local Similarity 88.0%; Pred. No. 2.4e-35;

Matches 183; Conservative 0; Mismatches 25; Indels 0; Gaps 0;





GenCore Version 5.1.3  
Copyright (c) 1993 2002 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 12, 2002, 16:50:46 ; Search time 636.023 Seconds

(without alignments) 9517.553 Million cell updates/sec

Title: US-09-697-123b-22

Perfect score: 208

Sequence: 1 tcaaggagaagcgtacgac.....ccggtggaaccgcagacat 208

Scoring table:

IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2054640 segs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenBank: \*  
1: gb\_ba: \*  
2: gb\_hlg: \*  
3: gb\_in: \*  
4: gb\_om: \*  
5: gb\_ov: \*  
6: gb\_pat: \*  
7: gb\_ph: \*  
8: gb\_pl: \*  
9: gb\_pr: \*  
10: gb\_ro: \*  
11: gb\_sts: \*  
12: gb\_sy: \*  
13: gb\_un: \*  
14: gb\_vl: \*  
15: em\_ba: \*  
16: em\_fun: \*  
17: em\_hum: \*  
18: em\_in: \*  
19: em\_mu: \*  
20: em\_om: \*  
21: em\_or: \*  
22: em\_ov: \*  
23: em\_pat: \*  
24: em\_ph: \*  
25: em\_pl: \*  
26: em\_ro: \*  
27: em\_sts: \*  
28: em\_un: \*  
29: em\_vl: \*  
30: em\_hlg\_hum: \*  
31: em\_hlg\_inv: \*  
32: em\_hlg\_other: \*  
33: em\_hlg\_mus: \*  
34: em\_hlg\_pln: \*  
35: em\_hlg\_rod: \*  
36: em\_hlg\_mam: \*  
37: em\_hlg\_vrt: \*  
38: em\_sy: \*  
39: em\_hlg\_hum: \*  
40: em\_hlg\_mus: \*  
41: em\_hlg\_other: \*

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	206.4	99.2	610	1	MTU318818
2	206.4	99.2	610	1	MTU318819
3	206.4	99.2	616	1	MTU318813
4	206.4	99.2	618	1	MTU318815
5	206.4	99.2	618	1	MTU318817
6	206.4	99.2	618	1	MTU318814
7	206.4	99.2	637	1	MTU318816
8	206.4	99.2	635	1	MTU318821
9	206.4	99.2	970	6	150706
10	206.4	99.2	3534	6	AX111339
11	206.4	99.2	3853	1	MTU12205
12	206.4	99.2	5084	1	MSGRPOB
13	206.4	99.2	19352	1	AEO06964
14	206.4	99.2	19770	1	MTC1376
15	147.2	70.8	3447	6	AR067447
16	147.2	70.8	3752	1	MSU24494
17	147.2	70.8	37617	1	MLB1790G
18	147.2	70.8	348950	1	MLEPRINT
19	98	47.1	3941	1	AF242549
20	88	42.3	3495	6	AX120631
21	88	42.3	328050	1	AP005275
22	88	42.3	349980	6	AX127144
23	80.4	38.7	32923	1	SCD82
24	45	21.6	15898	1	AX024393
25	44	21.6	15898	6	AX024286
26	44	21.2	7954	1	TA019223
27	43.4	20.9	134816	2	CNS08C9R
28	43.4	20.9	135378	2	CNS07CP9
29	42.6	20.5	1416	1	SREECAGEN
30	42	20.2	139467	1	D90905
31	41.8	20.1	34182	1	SC111
32	41.4	19.9	195859	14	AF281817
33	41	19.7	204050	1	AL646070
34	40.2	19.3	11103	1	AE004774
35	40	19.2	138792	2	AC113583
36	40	19.2	204050	1	AL646073
37	39.8	19.1	13431	1	AEO05044
38	39.8	19.1	119972	2	AP004029
39	39.6	19.0	598	8	AP000367
40	39.6	19.0	1271	1	PRE316161
41	39.6	19.0	1392	1	SVU04837
42	39.6	19.0	1392	1	PSERECAO
43	39.4	18.9	3056	3	PPU70864
44	39.4	18.9	5894	3	ACMHC
45	39.4	18.9	44109	1	SCSFI

#### ALIGNMENTS

RESULT 1  
LOCUS MTU318818 610 bp DNA circular BCT 09-AUG-2002  
DEFINITION Mycobacterium tuberculosis partial rpoB gene for RNA polymerase  
beta subunit, isolate 1415-97.  
ACCESSION AJ318818  
VERSION AJ318818.1 GI:22208412  
KEYWORDS RNA polymerase beta subunit; rpoB gene.  
SOURCE Mycobacterium tuberculosis.  
ORGANISM Mycobacterium tuberculosis  
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium  
tuberculosis complex.  
REFERENCE 1  
Herrera, L., Jimenez, M.S. and Saez, J.A.

Pred. No. is the number of results predicted by chance to have a

TITLE Molecular analysis of rifampin-resistant *Mycobacterium tuberculosis* isolated in Spain (1996-2001). Description of new alleles into rpoB gene and review

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 610)

AUTHORS Herrera, L.

JOURNAL Direct Submission

TITLE Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro Nacional Microbiologia, Ctra. Majadahonda-Pozuelo, Km 2.5, Majadahonda, Madrid 28220, SPAIN

FEATURES

source location/Qualifiers

1..610

/organism="Mycobacterium tuberculosis"

/isolate="1415-97"

/db\_xref="taxon:1773"

1..610

/gene="rpoB"

<1..>610

/gene="rpoB"

/codon\_start=1

/transl\_table=1

/product="RNA polymerase beta subunit"

/protein\_id="CAC87035.1"

/db\_xref="GI:22208413"

/translation="LDYRKLRPGEPTTESAQTLLNLFFEKRYDLARVGRVVK KLGLHVEPTTSTLTEDVVAITYEVLRLHSGOTTMTVPGVEVPEVETDIDHGNR RLRTVGLIQNIIRVGMSEMYRERMTQDVEAITPOTLINIRPVVAALKEFGTIS OPSOFMGNPNPLSGLFHKRLSLGPGGLSREBAGLEVRDVP"

BASE COUNT 122 a 191 c 202 g 95 t

ORIGIN

Query Match 99.28; Score 206.4; DB 1; Length 610;

Best Local Similarity 99.58; Pred. No. 1.5e-33;

Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TCAAGGAGAACGGCTAGACAGCTGGCCCGCGTGCCTATTAAGTCAACAAGAGCTCG 60

DB 80 TCAAGGAGAACGGCTAGACAGCTGGCCCGCGTGCCTATTAAGTCAACAAGAGCTCG 139

QY 61 GGCTGCATGTGGCGGAGCCCATCATCGTCGACGCTGACCGAAGAGAGCTGGGCCA 120

DB 140 GGCTGCATGTGGCGGAGCCCATCATCGTCGACGCTGACCGAAGAGAGCTGGGCCA 199

QY 121 CCATCGAATATCTGTCGCGCTTGACAGAGGTCAGACCAAGATGATCGTTCCGGCGGCG 180

DB 200 CCATCGAATATCTGTCGCGCTTGACAGAGGTCAGACCAAGATGATCGTTCCGGCGGCG 259

QY 181 TCGAGGTGCCGCTGGAACCGAGACAT 208

DB 260 TCGAGGTGCCGCTGGAACCGAGACAT 287

RESULT 2

LOCUS MTU318819 610 bp DNA circular BCT 09-AUG-2002

DEFINITION *Mycobacterium tuberculosis* partial rpoB gene for RNA polymerase beta subunit, isolate 1417-97.

ACCESSION AJ318819

VERSION AJ318819.1 GI:22208414

KEYWORDS RNA polymerase beta subunit; rpoB gene.

SOURCE *Mycobacterium tuberculosis*.

ORGANISM Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium tuberculosis complex.

REFERENCE 1

AUTHORS Herrera, L., Jimenez, M.S. and Saez, J.A.

TITLE Molecular analysis of rifampin-resistant *Mycobacterium tuberculosis* isolated in Spain (1996-2001). Description of new alleles into rpoB gene and review

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 610)

AUTHORS Herrera, L.

TITLE Direct Submission

JOURNAL Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro Nacional Microbiologia, Ctra. Majadahonda-Pozuelo, Km 2.5, Majadahonda, Madrid 28220, SPAIN

FEATURES

source location/Qualifiers

1..610

/organism="Mycobacterium tuberculosis"

/isolate="1417-97"

/db\_xref="taxon:1773"

1..610

/gene="rpoB"

<1..>610

/gene="rpoB"

/codon\_start=1

/transl\_table=1

/product="RNA polymerase beta subunit"

/protein\_id="CAC87036.1"

/db\_xref="GI:22208415"

/translation="LDYRKLRPGEPTTESAQTLLNLFFEKRYDLARVGRVVK KLGLHVEPTTSTLTEDVVAITYEVLRLHSGOTTMTVPGVEVPEVETDIDHGNR RLRTVGLIQNIIRVGMSEMYRERMTQDVEAITPOTLINIRPVVAALKEFGTIS OPSOFMGNPNPLSGLFHKRLSLGPGGLSREBAGLEVRDVP"

BASE COUNT 122 a 191 c 202 g 95 t

ORIGIN

Query Match 99.28; Score 206.4; DB 1; Length 610;

Best Local Similarity 99.58; Pred. No. 1.5e-33;

Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TCAAGGAGAACGGCTAGACAGCTGGCCCGCGTGCCTATTAAGTCAACAAGAGCTCG 60

DB 80 TCAAGGAGAACGGCTAGACAGCTGGCCCGCGTGCCTATTAAGTCAACAAGAGCTCG 139

QY 61 GGCTGCATGTGGCGGAGCCCATCATCGTCGACGCTGACCGAAGAGAGCTGGGCCA 120

DB 140 GGCTGCATGTGGCGGAGCCCATCATCGTCGACGCTGACCGAAGAGAGCTGGGCCA 199

QY 121 CCATCGAATATCTGTCGCGCTTGACAGAGGTCAGACCAAGATGATCGTTCCGGCGGCG 180

DB 200 CCATCGAATATCTGTCGCGCTTGACAGAGGTCAGACCAAGATGATCGTTCCGGCGGCG 259

QY 181 TCGAGGTGCCGCTGGAACCGAGACAT 208

DB 260 TCGAGGTGCCGCTGGAACCGAGACAT 287

RESULT 3

LOCUS MTU318813 616 bp DNA circular BCT 09-AUG-2002

DEFINITION *Mycobacterium tuberculosis* partial rpoB gene for RNA polymerase beta subunit, isolate 1763-97.

ACCESSION AJ318813

VERSION AJ318813.1 GI:22208402

KEYWORDS RNA polymerase beta subunit; rpoB gene.

SOURCE *Mycobacterium tuberculosis*.

ORGANISM Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium tuberculosis complex.

REFERENCE 1

AUTHORS Herrera, L., Jimenez, M.S. and Saez, J.A.

TITLE Molecular analysis of rifampin-resistant *Mycobacterium tuberculosis* isolated in Spain (1996-2001). Description of new alleles into rpoB gene and review

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 616)

AUTHORS Herrera, L.

JOURNAL Direct Submission

TITLE Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro Nacional Microbiologia, Ctra. Majadahonda-Pozuelo, Km 2.5, Majadahonda, Madrid 28220, SPAIN

FEATURES

source location/Qualifiers

1..616

gene  
CDS  
/organism="Mycobacterium tuberculosis"  
/isolate="1763-97"  
/db\_xref="taxon:1773"  
1..615  
/gene="rpoB"  
<1..>615  
/gene="rpoB"  
/codon\_start=1  
/transl\_table=11  
/product="RNA polymerase beta subunit"  
/protein\_id="CAC87030.1"  
/db\_xref="GI:22208403"

BASE COUNT 125 a 191 c 201 g 99 t  
ORIGIN

Query Match 99.2%; Score 206.4; DB 1; Length 616;  
Best Local Similarity 99.5%; Pred. No. 1.5e-33;  
Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TCAAGGAGAGCCGCTACGACCTGGCCGCGTGGCTATTAAGTCAACAAGAGCTCG 60  
|  
Db 80 TCAAGGAGAGCCGCTACGACCTGGCCGCGTGGCTATTAAGTCAACAAGAGCTCG 139  
Qy 61 GCGTCATGTCGGCGAGCCCATCAGCTCGTGCAGCTGACCGAAGAGCGTGGGCCA 120  
|  
Db 140 GCGTCATGTCGGCGAGCCCATCAGCTCGTGCAGCTGACCGAAGAGCGTGGGCCA 199  
Qy 121 CCATCGAATATCTGTGTCGCGCTTGACAGAGGTCAACACGATGATCGTCCGGGCGCG 180  
|  
Db 200 CCATCGAATATCTGTGTCGCGCTTGACAGAGGTCAACACGATGATCGTCCGGGCGCG 259  
Qy 181 TCGAGGTGCGCGGTGGAACCGACGACAT 208  
|  
Db 260 TCGAGGTGCGCGGTGGAACCGACGACAT 287

RESULT 4  
MTU318815 618 bp DNA circular BCT 09-AUG-2002  
LOCUS  
DEFINITION  
Mycobacterium tuberculosis partial rpoB gene for RNA polymerase  
beta subunit, isolate 2540-97.  
ACCESSION  
AJ318815.1 GI:22208406  
VERSION  
RNA polymerase beta subunit; rpoB gene.  
KEYWORDS  
Mycobacterium tuberculosis.  
SOURCE  
Mycobacterium tuberculosis  
ORGANISM  
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium  
tuberculosis complex.

REFERENCE  
AUTHORS  
TITLE  
Herrera, L., Jimenez, M. S. and Saez, J. A.  
Molecular analysis of rifampin-resistant Mycobacterium tuberculosis  
isolated in Spain (1996-2001). Description of new alleles into rpoB  
gene and review  
Unpublished  
2 (bases 1 to 618)  
Herrera, L.

JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
Direct Submission  
Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro  
Nacional Microbiologia, Ctra. Majadahonda-Pozuelo, Km 2.5,  
Majadahonda, Madrid, 28220, SPAIN

FEATURES  
SOURCE  
Location/Qualifiers

1..618  
/organism="Mycobacterium tuberculosis"  
/isolate="2540-97"  
/db\_xref="taxon:1773"  
1..618  
/gene="rpoB"  
<1..>618

gene  
CDS  
/gene="rpoB"  
/codon\_start=1  
/transl\_table=11  
/product="RNA polymerase beta subunit"  
/protein\_id="CAC87032.1"  
/db\_xref="GI:22208407"  
/transl\_table=11  
/transl\_table=11  
/product="RNA polymerase beta subunit"  
/protein\_id="CAC87034.1"  
/db\_xref="GI:22208411"

BASE COUNT 124 a 190 c 207 g 97 t  
ORIGIN

Query Match 99.2%; Score 206.4; DB 1; Length 616;  
Best Local Similarity 99.5%; Pred. No. 1.5e-33;  
Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TCAAGGAGAGCCGCTACGACCTGGCCGCGTGGCTATTAAGTCAACAAGAGCTCG 60  
|  
Db 95 TCAAGGAGAGCCGCTACGACCTGGCCGCGTGGCTATTAAGTCAACAAGAGCTCG 154  
Qy 61 GCGTCATGTCGGCGAGCCCATCAGCTCGTGCAGCTGACCGAAGAGCGTGGGCCA 120  
|  
Db 155 GCGTCATGTCGGCGAGCCCATCAGCTCGTGCAGCTGACCGAAGAGCGTGGGCCA 214  
Qy 121 CCATCGAATATCTGTGTCGCGCTTGACAGAGGTCAACACGATGATCGTCCGGGCGCG 180  
|  
Db 215 CCATCGAATATCTGTGTCGCGCTTGACAGAGGTCAACACGATGATCGTCCGGGCGCG 274  
Qy 181 TCGAGGTGCGCGGTGGAACCGACGACAT 208  
|  
Db 275 TCGAGGTGCGCGGTGGAACCGACGACAT 302

RESULT 5  
MTU318817 618 bp DNA circular BCT 09-AUG-2002  
LOCUS  
DEFINITION  
Mycobacterium tuberculosis partial rpoB gene for RNA polymerase  
beta subunit, isolate 2348-98.  
ACCESSION  
AJ318817.1 GI:22208410  
VERSION  
RNA polymerase beta subunit; rpoB gene.  
KEYWORDS  
Mycobacterium tuberculosis.  
SOURCE  
Mycobacterium tuberculosis  
ORGANISM  
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium  
tuberculosis complex.

REFERENCE  
AUTHORS  
TITLE  
Herrera, L., Jimenez, M. S. and Saez, J. A.  
Molecular analysis of rifampin-resistant Mycobacterium tuberculosis  
isolated in Spain (1996-2001). Description of new alleles into rpoB  
gene and review  
Unpublished  
2 (bases 1 to 618)  
Herrera, L.

JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
Direct Submission  
Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro  
Nacional Microbiologia, Ctra. Majadahonda-Pozuelo, Km 2.5,  
Majadahonda, Madrid, 28220, SPAIN

FEATURES  
SOURCE  
Location/Qualifiers

1..618  
/organism="Mycobacterium tuberculosis"  
/isolate="2348-98"  
/db\_xref="taxon:1773"  
1..618  
/gene="rpoB"  
<1..>618  
/gene="rpoB"  
/codon\_start=1  
/transl\_table=11  
/product="RNA polymerase beta subunit"  
/protein\_id="CAC87034.1"  
/db\_xref="GI:22208411"

/translation="TDEALLDIYRKLRPGEPPTKESAOITLLENLFFKERYDLARV  
YKVNKKLGLHVEPTSTLTEDVAVATEYLVRHLEGOTTMTVPGGVEVETDD  
HFGNRRLRTVCELIONQIRVGSRMERVRREMTTODVEAITPQTLINIRPVAAIK  
PFQTSQDSQSGMGNPNLSGLTHKRLSLAAGPGLSRRAGLEVRDV"  
BASE COUNT 124 a 192 c 207 g 95 t  
ORIGIN

Query Match 99.2%: Score 206.4; DB 1: Length 618;  
Best Local Similarity 99.5%: Pred. No. 1.5e-33;  
Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 TCAAGAGAGAGCGCTACGACCTGGCCCGCTCGCTCTAAGTCAACAGAACCTCG 60  
|||||  
Db 95 TCAAGAGAGAGCGCTACGACCTGGCCCGCTCGCTCTAAGTCAACAGAACCTCG 154  
Oy 61 GCGTCAATATCTGTGCTCGCTTCGACGAGGTCAGACCGATGATCGTTCCGGGGCGG 120  
|||||  
Db 155 GCGTCAATATCTGTGCTCGCTTCGACGAGGTCAGACCGATGATCGTTCCGGGGCGG 214  
Oy 121 CCATGCAATATCTGTGCTCGCTTCGACGAGGTCAGACCGATGATCGTTCCGGGGCGG 180  
|||||  
Db 215 CCATGCAATATCTGTGCTCGCTTCGACGAGGTCAGACCGATGATCGTTCCGGGGCGG 274  
Oy 181 TCGAGGTGCGCGGTGGAACCGACGACAT 208  
|||||  
Db 275 TCGAGGTGCGCGGTGGAACCGACGACAT 302

RESULT 6  
MTU318814 633 bp DNA circular BCT 09-AUG-2002  
LOCUS  
DEFINITION  
Mycobacterium tuberculosis partial rpoB gene for RNA polymerase  
beta subunit, isolate 1058-97.  
ACCESSION  
AJ318814  
VERSION  
AJ318814.1 GI:22208404  
KEYWORDS  
RNA polymerase beta subunit; rpoB gene.  
SOURCE  
Mycobacterium tuberculosis.  
ORGANISM  
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium  
tuberculosis complex.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
Herrera, L., Jimenez, M.S. and Saez, J.A.  
Molecular analysis of rifampin-resistant Mycobacterium tuberculosis  
isolated in Spain (1996-2001). Description of new alleles into rpoB  
gene and review  
Unpublished  
2 (bases 1 to 633)

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
Direct Submission  
Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro  
Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5,  
Majadahonda. Madrid. 28220, SPAIN  
Location/Qualifiers  
1. 633  
/organism="Mycobacterium tuberculosis"  
/isolate="1058-97"  
/db\_xref="taxon:1773"  
1. 633  
/gene="rpoB"  
/codon\_start=1  
/transl\_table=1  
/product="RNA polymerase beta subunit"  
/protein\_id="CAC87033.1"  
/db\_xref="GI:22208405"  
/translation="DNTVCTDEALLDIYRKLRPGEPPTKESAOITLLENLFFKERYDL  
ARVGRYKKNKGLHVEPTSTLTEDVAVATEYLVRHLEGOTTMTVPGGVEVETDD  
TDDHFGNRRLRTVCELIONQIRVGSRMERVRREMTTODVEAITPQTLINIRPVV  
AAIKPFQTSQDSQSGMGNPNLSGLTHKRLSLAAGPGLSRRAGLEVRDV"

gene  
CDS

BASE COUNT  
ORIGIN

Query Match 99.2%: Score 206.4; DB 1: Length 633;  
Best Local Similarity 99.5%: Pred. No. 1.5e-33;  
Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 TCAAGAGAGAGCGCTACGACCTGGCCCGCTCGCTCTAAGTCAACAGAACCTCG 60  
|||||  
Db 110 TCAAGAGAGAGCGCTACGACCTGGCCCGCTCGCTCTAAGTCAACAGAACCTCG 169  
Oy 61 GCGTCAATATCTGTGCTCGCTTCGACGAGGTCAGACCGATGATCGTTCCGGGGCGG 120  
|||||  
Db 170 GCGTCAATATCTGTGCTCGCTTCGACGAGGTCAGACCGATGATCGTTCCGGGGCGG 229  
Oy 121 CCATGCAATATCTGTGCTCGCTTCGACGAGGTCAGACCGATGATCGTTCCGGGGCGG 180  
|||||  
Db 230 CCATGCAATATCTGTGCTCGCTTCGACGAGGTCAGACCGATGATCGTTCCGGGGCGG 289  
Oy 181 TCGAGGTGCGCGGTGGAACCGACGACAT 208  
|||||  
Db 290 TCGAGGTGCGCGGTGGAACCGACGACAT 317

RESULT 7  
MTU318816 637 bp DNA circular BCT 09-AUG-2002  
LOCUS  
DEFINITION  
Mycobacterium tuberculosis partial rpoB gene for RNA polymerase  
beta subunit, isolate 1255-98.  
ACCESSION  
AJ318816  
VERSION  
AJ318816.1 GI:22208408  
KEYWORDS  
RNA polymerase beta subunit; rpoB gene.  
SOURCE  
Mycobacterium tuberculosis.  
ORGANISM  
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium  
tuberculosis complex.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
Herrera, L., Jimenez, M.S. and Saez, J.A.  
Molecular analysis of rifampin-resistant Mycobacterium tuberculosis  
isolated in Spain (1996-2001). Description of new alleles into rpoB  
gene and review  
Unpublished  
2 (bases 1 to 637)

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
Direct Submission  
Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro  
Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5,  
Majadahonda. Madrid. 28220, SPAIN  
Location/Qualifiers  
1. 637  
/organism="Mycobacterium tuberculosis"  
/isolate="1255-98"  
/db\_xref="taxon:1773"  
1. 637  
/gene="rpoB"  
/codon\_start=1  
/transl\_table=1  
/product="RNA polymerase beta subunit"  
/protein\_id="CAC87033.1"  
/db\_xref="GI:22208409"  
/translation="VCTDEALLDIYRKLRPGEPPTKESAOITLLENLFFKERYDLARV  
GRYVKNKGLHVEPTSTLTEDVAVATEYLVRHLEGOTTMTVPGGVEVETDD  
IDHFGNRRLRTVCELIONQIRVGSRMERVRREMTTODVEAITPQTLINIRPVVAAI  
KEFFQTSQDSQSGMGNPNLSGLTHKRLSLAAGPGLSRRAGLEVRDVHPSH"

gene  
CDS

BASE COUNT  
ORIGIN

Query Match 99.2%: Score 206.4; DB 1: Length 637;  
Best Local Similarity 99.5%: Pred. No. 1.5e-33;  
Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Oy 1 TCAAGAGAGAGCGCTACGACCTGGCCCGCTCGCTCTAAGTCAACAGAACCTCG 60



|||||  
Db 101 TCAAGAGAAAGCGCTACAGACCTGCGCCGCTGCGCTATTAAGGTCAAGAAGAGCTCG 160  
OY 61 GGCTGCATGTCGGCGAGCCCATCAGCTGTGACGCTTACCGAAGAACAGCTCGTGGCCA 120  
Db 161 GGCTGCATGTCGGCGAGCCCATCAGCTGTGACGCTTACCGAAGAACAGCTCGTGGCCA 220  
OY 121 CCATCGAATATCTGCTGCGCTTGCAGAGGGTCAGACGACGATGATCGTTCCGGGGGGCG 180  
Db 221 CCATCGAATATCTGCTGCGCTTGCAGAGGGTCAGACGACGATGATCGTTCCGGGGGGCG 280  
OY 181 TCGAGGTCCCGGTGGAACCGACGACAT 208  
Db 281 TCGAGGTCCCGGTGGAACCGACGACAT 308  
RESULT 8  
LOCUS MTU318821 639 bp DNA circular BCT 09-AUG-2002  
DEFINITION Mycobacterium tuberculosis partial rpoB gene for RNA polymerase  
beta subunit, isolate 1071-98.  
ACCESSION AJ318821  
VERSION AJ318821.1 GI:22208418  
KEYWORDS RNA polymerase beta subunit; rpoB gene.  
SOURCE Mycobacterium tuberculosis.  
ORGANISM Mycobacterium tuberculosis  
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
Corynebacteriineae; Mycobacteriaceae; Mycobacterium; Mycobacterium  
tuberculosis complex.  
REFERENCE 1  
AUTHORS Herrera, L., Jimenez, M. S. and Saez, J. A.  
TITLE Molecular analysis of rifampin-resistant Mycobacterium tuberculosis  
isolated in Spain (1996-2001). Description of new alleles into rpoB  
gene and review  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 639)  
AUTHORS Herrera, L.  
TITLE Direct Submission  
JOURNAL Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro  
Nacional Microbiologia, Ctra. Majadahonda-Pozuelo, Km 2.5,  
Majadahonda, Madrid 28220, SPAIN  
FEATURES  
source  
1. .639  
/organism="Mycobacterium tuberculosis"  
/isolate="1071-98"  
/db\_xref="taxon:1773"  
1. .639  
/gene="rpoB"  
<1. .>639  
/gene="rpoB"  
/codon\_start=1  
/transl\_table=11  
/product="RNA polymerase beta subunit"  
/protein\_id="CAC87038.1"  
/db\_xref="GI:22208418"  
/translation="RRTDALDLIYRKLRRGPPPTKESQTLLENLFKKRYDLAR  
GRYKVKKLLHVGEPITSLTEEDVAVTIEVLVHLHGQTMVPGVGEVERD  
IDHFGNRRLTVELIIONIRVMSRMERVRERMTQOVELIPTLLINIPAVAAI  
KEPFGTSLSOFWVONNPLSGILTKRRLSALRPGGLSRERAGLEVNDVPHSHS"  
BASE COUNT 126 a 202 c 212 g 99 t  
ORIGIN  
Query Match 99.2%; Score 206.4; DB 1; Length 639;  
Best Local Similarity 99.5%; Pred. No. 1.5e-33;  
Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
OY 1 TCAAGAGAAAGCGCTACAGACCTGCGCCGCTGCGCTATAAGGTCAACAGACCTCG 60  
Db 101 TCAAGAGAAAGCGCTACAGACCTGCGCCGCTGCGCTATAAGGTCAACAGACCTCG 160  
OY 61 GGCTGCATGTCGGCGAGCCCATCAGCTGTGACGCTTACCGAAGAACAGCTCGTGGCCA 120  
Db 161 GGCTGCATGTCGGCGAGCCCATCAGCTGTGACGCTTACCGAAGAACAGCTCGTGGCCA 220

OY 121 CCATCGAATATCTGCTGCGCTTGCAGAGGGTCAGACGACGATGATCGTTCCGGGGGGCG 180  
Db 221 CCATCGAATATCTGCTGCGCTTGCAGAGGGTCAGACGACGATGATCGTTCCGGGGGGCG 280  
OY 181 TCGAGGTCCCGGTGGAACCGACGACAT 208  
Db 281 TCGAGGTCCCGGTGGAACCGACGACAT 308  
RESULT 9  
LOCUS 150706 970 bp DNA linear PAT 07-OCT-1997  
DEFINITION Sequence 1 from patent US 5643723.  
ACCESSION 150706  
VERSION 150706.1 GI:2472409  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 970)  
AUTHORS Persing, D. H., Hunt, J. J., Young, K. K. Y., Felmlie, T. A., Roberts, G. D.  
and Whelan, A. Christian.  
TITLE Detection of a genetic locus encoding resistance to rifampin in  
mycobacterial cultures and in clinical specimens  
JOURNAL Patent: US 5643723-A I 01-JUL-1997;  
FEATURES  
source  
1. .970  
/organism="unknown"  
BASE COUNT 182 a 302 c 330 g 156 t  
ORIGIN  
Query Match 99.2%; Score 206.4; DB 6; Length 970;  
Best Local Similarity 99.5%; Pred. No. 1.4e-33;  
Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
OY 1 TCAAGAGAAAGCGCTACAGACCTGCGCCGCTGCGCTATAAGGTCAACAGACCTCG 60  
Db 26 TCAAGAGAAAGCGCTACAGACCTGCGCCGCTGCGCTATAAGGTCAACAGACCTCG 85  
OY 61 GGCTGCATGTCGGCGAGCCCATCAGCTGTGACGCTTACCGAAGAACAGCTCGTGGCCA 120  
Db 86 GGCTGCATGTCGGCGAGCCCATCAGCTGTGACGCTTACCGAAGAACAGCTCGTGGCCA 145  
OY 121 CCATCGAATATCTGCTGCGCTTGCAGAGGGTCAGACGACGATGATCGTTCCGGGGGGCG 180  
Db 146 CCATCGAATATCTGCTGCGCTTGCAGAGGGTCAGACGACGATGATCGTTCCGGGGGGCG 205  
OY 181 TCGAGGTCCCGGTGGAACCGACGACAT 208  
Db 206 TCGAGGTCCCGGTGGAACCGACGACAT 233  
RESULT 10  
LOCUS AX111339 3534 bp DNA linear PAT 30-APR-2001  
DEFINITION Sequence 2072 from Patent WO0123604.  
ACCESSION AX111339  
VERSION AX111339.1 GI:13927631  
KEYWORDS  
SOURCE Mycobacterium tuberculosis.  
ORGANISM Mycobacterium tuberculosis  
Bacteria; Filumicutes; Actinobacteria; Actinobacteridae;  
Actinomycetales; Corynebacteriineae; Mycobacteriaceae;  
Mycobacterium; Mycobacterium tuberculosis complex.  
REFERENCE 1 (bases 1 to 3534)  
AUTHORS Bergeron, M. G., Boissinot, M., Huelstsky, A., m Nard, C., Ouellette, M.,  
Picard, F. J. and Roy, P. H.  
TITLE Highly conserved genes and their use to generate probes and primers  
for detection of microorganisms  
JOURNAL Patent: WO 0123604-A 2072 05-APR-2001;  
FEATURES  
Infectio Diagnostic (I.D.I.) INC. (CA)  
Location/Qualifiers

```
source 1..3534
/organism="Mycobacterium tuberculosis"
/strain="Rv"
/db_xref="taxon:1773"
BASE COUNT 679 a 1081 c 1188 g 586 t
ORIGIN
Query Match 99.2%; Score 206.4; DB 6; Length 3534;
Best Local Similarity 99.5%; Pred. No. 1.3e-33;
Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TCACGAGAACGCTTACGACCTGCCCCGCGTGGTCTATTAAGTCAACAAGAACCTCG 60
DB 902 TCACGAGAACGCTTACGACCTGCCCCGCGTGGTCTATTAAGTCAACAAGAACCTCG 961
OY 61 GCGTCATGTCGCGGAGCCCATCAGCTGCGTGCAGCTGACCCGAAGAAGCGTCGGCCCA 120
DB 962 GCGTCATGTCGCGGAGCCCATCAGCTGCGTGCAGCTGACCCGAAGAAGCGTCGGCCCA 1021
OY 121 CCATCGAATATCTGTCGCTTGCACGAGGTCGACGACGATGATCTTCCGGCGCGC 180
DB 1022 CCATCGAATATCTGTCGCTTGCACGAGGTCGACGACGATGATCTTCCGGCGCGC 1081
OY 181 TCAGAGTCCGCGTGGAAACCGACGACAT 208
DB 1082 TCAGAGTCCGCGTGGAAACCGACGACAT 1109

RESULT 11
MTU12205 3853 bp DNA linear BCT 02-MAR-2000
LOCUS Mycobacterium tuberculosis H37Rv RNA-polymerase beta subunit (rpoB)
DEFINITION gene, partial cds.
ACCESSION U12205
VERSION U12205.1 GI:515684
KEYWORDS
SOURCE Mycobacterium tuberculosis.
ORGANISM Mycobacterium tuberculosis
Bacteria; Firmicutes; Actinobacteria; Actinomycetales;
Actinomycetales; Corynebacteriaceae; Mycobacteriaceae;
Mycobacterium; Mycobacterium tuberculosis complex.
1 (bases 1 to 3853)
Imboden,P., Toller,R., Marchesi,F., Telenti,A., Bodmer,T.,
Cole,S., Schopfer,K. and Burkart,T.
The rpoB gene of Mycobacterium tuberculosis
Unpublished
2 (bases 1 to 3853)
Imboden,P.
Submitted (11-JUL-1994) Paul Imboden, Institute for Medical
Microbiology, University of Berne, Friedbuehlstrasse 51, Berne,
3010, Switzerland
Location/Qualifiers
1..3853
/organism="Mycobacterium tuberculosis"
/strain="H37Rv"
/db_xref="taxon:1773"
576..>3853
/gene="rpoB"
576..>3853
/gene="rpoB"
/codon_start=1
/transl_table=11
/product="RNA-polymerase beta subunit"
/protein_id="AAA20242.2"
/db_xref="GI:7144499"
/translation="MLEGCIADRSKTAASPSRPOSSNNNSVPCAPNRVSFAKL
REPLEVPLDVOYDSEBMLIGSPMRSSAERDQVNPVGLLEVLTELSTEDFSGS
MSLSDFRFDVAPVDECKDKMTTAAPLFTVAEFLNNNTGEIKSQVEMGDFPM
TEKGTFFIINGTERVVSQVRSQVETIDKSTDKLHVSQVPSGAMLEFVDK
RDYGVIRIDRKRPQVTVLKLAKMTSEQIVERSGFSIMRSTLEKNTVOTDEALD
IYRLRPEPTKESATLLENLFPEKRYDLAVGRYKVKKILGLHVGEPITSSITL
EDVAVATIEYLRLHEGOTMTVPGGVEVPEPTDIDHFRRLRTVGLDELQNIIRVG
```

```
MSMRERVAREMTTQDEATTPOTLINIRPVNAIKEPFTSOLSOFPMDNNPLSGLT
HKRLSLGEGCISREAGLEVDYRSHRGRCPIETEPGPITGLISLTVARNP
FGTETPRKRVGVSDGIVTLADEEDHVVQAQNSPLDADGRVERVLRKAG
EVERVPSSEVDYMDVSRQWVAATAMIPELHDDANRALMGAMROQAVPLRSEAP
LVGGMELRAAIDAATSSQSESGVIEVSADYITVMDHNDRTYRMRFAHSNHTC
ANOCPIVDAGDRVACGVADGCTDGEALCNLLVAIMPEGHNYBDAILISRL
VEEDVLSIHIEHEIDARDTKLGAETIDKINISDEVADLDERGIVRIGAEVSD
DILGVKTPKGTETPEERLLAIGERAREPRDSIKVPHRESKVRIGIRFNSHD
EDELPRACVNELVYVVAQKRISDGBKLGRHGNKVIGITLPEVDMPLADGTPVDI
ILNTHGVPRRMNIGQILETHLGMCAHSGMKVDAKVPMAARLPDELLEAQPNAYS
TPVFDGAQEAELQGLLSCITLPNBDGVLVADADKAMLFDRSGSEPPPEVTVGYIM
KLHLVYDVKIHAIRSTGPYSMTQOPLGKAKOFGQRFREMECAWQAVCAATVLOGLL
TIKS"
BASE COUNT 723 a 1173 c 1293 g 664 t
ORIGIN
Query Match 99.2%; Score 206.4; DB 1; Length 3853;
Best Local Similarity 99.5%; Pred. No. 1.3e-33;
Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TCACGAGAACGCTTACGACCTGCCCCGCGTGGTCTATTAAGTCAACAAGAACCTCG 60
DB 1477 TCACGAGAACGCTTACGACCTGCCCCGCGTGGTCTATTAAGTCAACAAGAACCTCG 1536
OY 61 GCGTCATGTCGCGGAGCCCATCAGCTGCGTGCAGCTGACCCGAAGAAGCGTCGGCCCA 120
DB 1537 GCGTCATGTCGCGGAGCCCATCAGCTGCGTGCAGCTGACCCGAAGAAGCGTCGGCCCA 1596
OY 121 CCATCGAATATCTGTCGCTTGCACGAGGTCGACGACGATGATCTTCCGGCGCGC 180
DB 1597 CCATCGAATATCTGTCGCTTGCACGAGGTCGACGACGATGATCTTCCGGCGCGC 1656
OY 181 TCAGAGTCCGCGTGGAAACCGACGACAT 208
DB 1657 TCAGAGTCCGCGTGGAAACCGACGACAT 1684

RESULT 12
MSGRPOB 5084 bp DNA linear BCT 13-SEP-1994
LOCUS Mycobacterium tuberculosis RNA polymerase beta-subunit (rpoB)
DEFINITION gene, complete cds and RNA polymerase beta'-subunit rpoC gene,
partial cds.
ACCESSION L27989
VERSION L27989.1 GI:466333
KEYWORDS RNA polymerase beta-subunit; rpoB gene.
SOURCE Mycobacterium tuberculosis (strain Rv) DNA.
ORGANISM Mycobacterium tuberculosis
Bacteria; Actinobacteria; Actinomycetales;
Corynebacteriaceae; Mycobacteriaceae; Mycobacterium;
Mycobacterium tuberculosis complex.
1 (bases 1 to 5084)
Miller,L.P., Crawford,J.T. and Shinnick,T.M.
The rpoB gene of Mycobacterium tuberculosis
Antimicrob. Agents Chemother. 38 (4), 805-811 (1994)
Location/Qualifiers
1..5084
/organism="Mycobacterium tuberculosis"
/strain="Rv"
/db_xref="taxon:1773"
1065..4598
/gene="rpoB"
1065..4598
/gene="rpoB"
/codon_start=1
/evidence=experimental
/transl_table=11
/product="RNA polymerase beta-subunit"
/protein_id="AAA21416.1"
/db_xref="GI:466334"
/translation="MLEGCIADRSKTAASPSRPOSSNNNSVPCAPNRVSFAKL
```

REPLEVGLLDVOTDSFEMLIGSPRMRESAERGDVNPVGCLEVLIELSFIEDFSS  
MSLSFSDPRFDVAKAPVDECKDKMTIAPLEPVIAKEFINNTGKIKQYFMGDFPM  
TEKGFITTINGERKVVSVLSVGVYFDETDIDSKDTLHNSGKIPSRGAMLEFDVAK  
RDTVGRIIDRRKRPQVTLKALAGMTSEQIYERGFSEIMSTLEKNVTGDELDALD  
IYRLKRPPEPTKESAOTLLENLFKEKRYDLARGRKXKKTGLGNGVPTISLTD  
EEDVATIEYLVRHLEGOTMTVPGVGEVPEETDIDHFGRRRLRTYGELOQNIORG  
MSRMVEREMRTODVEAITPOTLINIRPVVAIKERFETQSLSQSDMDONPLSGLT  
HRRRLSALGCGLSRERAGLEVDRVPHSHYRMRKPIETPEBPNIGLIGSLSVARVP  
RGEFIEPTIRKRVYDVSVDEIYLTADDEDRHVQAQNSPIDADGEFVPLVRKAG  
EVEYVSSVVDVMSPRQWVSATAMIPLEHDANRSLAMQAOAVPLVRSKAP  
LVGTGAGELKAAIDAATSSQSGVIEEVSADYITVMHNGRRTYRMRKPARSHSTL  
ANOCPIVDAGDRVEAGVYIADGPTDDEMLGKNLVATIMPGEHNVEDALISNLT  
VEEDVATIEYLVRHLEGOTMTVPGVGEVPEETDIDHFGRRRLRTYGELOQNIORG  
DILVGTMPKGETELTPEERILRAIFGKADEVDTSLKPHGSGKVIGIRVFSRED  
EDELPAQNELVRYVVAOKRKISDGDLAGHNGKGVIGILVPEMDPLADGTPVDI  
ILNTHGVPRRMNIGIOLETHLGCASHSGMKVDAAKGVDMARLPDELLAHANAYTS  
TVFVGDAGEAEIOLGSLCTLPNRDGDVYVADGKAMLPDGRSGEPFYPVTVYVMIM  
KLHLHLYDVKIHARSTGYSMTIQPLGKAQFGORFGEBCNMAQVGAAYTLOEL  
TIKSDYTVGRVYVEAIVKGENIPEPGIPESFKVLKELOSLCLNVEVLSDDAATIEL  
RGEDEDELERAANGLINLSNESASVEDLA"

gene  
4641..5084  
/gene="tpoc"  
/db\_xref="GI:1380217.1"

CDS  
4641..5084  
/gene="tpoc"  
/codon\_start=1  
/transl\_table=1  
/product="RNA polymerase beta -subunit"  
/protein\_id="AAK4147.1"  
/db\_xref="GI:1380217.1"

BASE COUNT 969 a 1534 c 1691 g 890 t

Query Match 99.2% Score 206.4 DB 1 Length 5084  
Best Local Similarity 99.5% Pred. No. 1.3e-33  
Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 TCAAGAGAACGCGTACACACCGCGCGCGGTACGCGTAAAGCTACACAAAGAGCTCG 60  
1966 TCAAGAGAACGCGTACACACCGCGCGCGGTACGCGTAAAGCTACACAAAGAGCTCG 2025  
61 GCGTCATGTCGCGGAGCGCCATCAACGTCGTCGACGCGAAGAGAGCTCGGCCA 120  
2026 GCGTCATGTCGCGGAGCGCCATCAACGTCGTCGACGCGAAGAGAGCTCGGCCA 2085  
121 CCATCAATATCTGGTCCGCTTGACAGAGGTCAGACGACGATGTCGCGGCGCGC 180  
2086 CCATCAATATCTGGTCCGCTTGACAGAGGTCAGACGACGATGTCGCGGCGCGC 2145  
181 TCGAGGTCGCGGTGGAACCGACGACAT 208  
2146 TCGAGGTCGCGGTGGAACCGACGACAT 2173

RESULT 13  
AE006964 19352 bp DNA linear BCF 27-APR-2001  
LOCUS Mycobacterium tuberculosis CDC1551, section 50 of 280 of the  
DEFINITION complete genome.  
ACCESSION AE006964 AE000516  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Mycobacterium tuberculosis CDC1551.  
Mycobacterium tuberculosis CDC1551.  
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
Actinomycetales; Corynebacteriaceae; Mycobacteriaceae;  
Mycobacterium; Mycobacterium tuberculosis complex.  
1 (bases 1 to 19352)  
Fleischmann, R.D., Alland, D., Eisen, J.A., Carpenter, L., White, O.,  
Peterson, J., DeBoy, R., Dodson, R., Gwinn, M., Haft, D., Hickey, E.,  
Kolony, J.F., Nelson, W.C., Umayam, L.A., Ermolaeva, M.,

TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
Unpublished  
2 (bases 1 to 19352)  
Fleischmann, R.D., Alland, D., Eisen, J.A., Carpenter, L., White, O.,  
Peterson, J., DeBoy, R., Dodson, R., Gwinn, M., Haft, D., Hickey, E.,  
Kolony, J.F., Nelson, W.C., Umayam, L.A., Ermolaeva, M.,  
Salzberg, S.L., Delcher, A., Uitterback, T., Weidman, J., Khouri, H.,  
Gill, J., Mikula, A. and Bishai, W.  
Direct Submission  
Submitted (25-Apr-2001) The Institute for Genomic Research, 9712  
Medical Center Dr., Rockville, MD 20850, USA  
Location/Qualifiers  
1..19352  
/organism="Mycobacterium tuberculosis CDC1551"  
/strain="CDC1551"  
/db\_xref="taxon:83331"  
/note="clinical strain"  
163..3699  
/gene="MT0695"  
163..3699  
/note="MT0695"  
/note="similar to GB:L27989 GB:I05910 GB:U12205 SP:P47766  
PID:149992; identified by sequence similarity; putative"  
/codon\_start=1  
/transl\_table=1  
/product="DNA-directed RNA polymerase, beta subunit"  
/protein\_id="AAK4921.1"  
/db\_xref="GI:1380218"

gene  
CDS

gene  
CDS

Salzberg, S.L., Delcher, A., Uitterback, T., Weidman, J., Khouri, H.,  
Gill, J., Mikula, A. and Bishai, W.  
Whole genome comparison of Mycobacterium tuberculosis clinical and  
laboratory strains  
Unpublished  
2 (bases 1 to 19352)  
Fleischmann, R.D., Alland, D., Eisen, J.A., Carpenter, L., White, O.,  
Peterson, J., DeBoy, R., Dodson, R., Gwinn, M., Haft, D., Hickey, E.,  
Kolony, J.F., Nelson, W.C., Umayam, L.A., Ermolaeva, M.,  
Salzberg, S.L., Delcher, A., Uitterback, T., Weidman, J., Khouri, H.,  
Gill, J., Mikula, A. and Bishai, W.  
Direct Submission  
Submitted (25-Apr-2001) The Institute for Genomic Research, 9712  
Medical Center Dr., Rockville, MD 20850, USA  
Location/Qualifiers  
1..19352  
/organism="Mycobacterium tuberculosis CDC1551"  
/strain="CDC1551"  
/db\_xref="taxon:83331"  
/note="clinical strain"  
163..3699  
/gene="MT0695"  
163..3699  
/note="MT0695"  
/note="similar to SP:P37871; identified by sequence  
similarity; putative"  
/codon\_start=1  
/transl\_table=1  
/product="DNA-directed RNA polymerase, beta-prime subunit"  
/protein\_id="AAK44922.1"  
/db\_xref="GI:1380219"

SPASGRPLAMPRLDMVTLGYLTTEVPEDTGEYOPASGDHPETGVYSSPAEAIAMADR  
GVLVSRAKIVRLTLQRLPVEIEALFEGHSGMOGDAMMAETLIGRMNELILPGP  
FVKMKHKKVOAAIINDIARERPMIVAOCTVDKACGPMWARSQVYTSMADEVLP  
RKKEILLDEERADVKYERKORCALNHNDENEALVEIMKESATDEVCOALREPDNP  
ITITVBSGATGNTOTRTTLGAKMGLTVNKGETIPRPVSSFEFRELTVLEYEINTGA  
RKGLADALTADSGYTLRLVDSQDIVREHDCOTERQIVELAEAPDGLIIDP  
YIETSAVARTGLDAVEAGNVIVERGODIGDEIDALLAAGITLYKRVSVLCAST  
GVCATCGSRMGAKIADVIGEAAGIIVAAOSIGEPGQITMRTPHOGVVEDITGCLPR  
VOELFEARVPKAPLADVYGRVLEDEGERFYKITIYPDGGCEVYVYDKISKRQRLR  
FKHEDSERVLSGDHVEVGOOCEGSADEPLVRQGREVOIHLVREVOEYRAOG  
VSTHDKHIEYIVKOMLRVYITIDSGTEFLPGSLIDRAEFEMENRRVAVAEGLPAGR  
PVLAKITKASLATDSWLSASFQETTRVLTDAAINCRSDKLNGLKENYIIGKILPGT  
GIMRRIIAVQPEPEARAAYTIPSYEDDYSPDFGAATGAAPLDDYGYSDYR"  
/complement(.7691. .8065)  
/gene="MT0697"  
/complement(.7691. .8065)  
/gene="MT0697"  
/note="Identified by glimmer2; putative"  
/codon\_start=1  
/transl\_table=11  
/product="hypothetical protein"  
/protein\_id="AAK44923.1"  
/db\_xref="GI:13880220"  
/translation="MPSAAATINPCNANASAMERSGLLECVAAGLDEDPGFEFTADKL  
NPDGSSRRVPRROADGIAITHVERGGQKSGGQGVVPORMHGFPPALAMODRLIH  
GEOTQNRIDQAEVRFVCSPT"  
/complement(.8058. .9972)  
/gene="MT0698"  
/note="This region contains an authentic frame shift and  
is not the result of a sequencing artifact; identified by  
glimmer2; putative; conserved hypothetical protein,  
authentic frameshift"  
10167. .10925  
/gene="MT0699"  
10167. .10925  
/gene="MT0699"  
/note="Identified by match to PRM protein family HM  
PF01261"  
/codon\_start=1  
/transl\_table=11  
/product="AP endonuclease, family 2"  
/protein\_id="AAK44924.1"  
/db\_xref="GI:13880221"  
/translation="MLIGSHVSPDPIAAAEAGADVQIFIGNQSKAKPRDDAA  
ALKATIPPIYHAPYLINLASNNRYRIPSRKITDETCAADIGAAAVIHGGVAD  
DNDIDKGRMRKALDRLETEVYLTENTAGSDHMAKRFDTIARLMDVIDTGGFC  
LDTCHTWAGEALTDVDRIKAITGRIDLVHCNDRDEAGSGRORHAMVSGQIDPDL  
LVAAVKAAGAPVICTADQGRKDIATFLKERTGS"  
10957. .11799  
/gene="MT0700"  
10957. .11799  
/gene="MT0700"  
/note="similar to GB:U00012 PID:466863; identified by  
sequence similarity; putative"  
/codon\_start=1  
/transl\_table=11  
/product="hydrolase/esterase, putative"  
/protein\_id="AAK44925.1"  
/db\_xref="GI:13880222"  
/translation="MLRVAAILLAAVLAFAAGCSGTRLAAGFGNNSVHTLVDVAGR  
SYRLKPYGLPSAPLVVMLHGFSGSAQASRSQADELADSEKFLVAVPGYHRAN  
ANGGCCGRPARREGVDIGFRAVAVDIAANNVSDIPARYVITGMSNGAIVSYTACNT  
SIFAAIGVSGOLDPCOSPRPVSYIHHGADADLVHRHGGAGPARIDCPYPDLN  
AFWRVNRGALDITTBGPTTSATCADNRRVYLLIYDDAGHRMPSATOTLWRFPA  
AHRP"  
11859. .13487  
/gene="MT0701"  
11859. .13487  
/gene="MT0701"  
/note="similar to SP:P33224 GB:L20915 PID:457172  
PID:457174 PID:537028; identified by sequence similarity;  
putative"  
/codon\_start=1

/transl\_table=11  
/product="acyl-CoA dehydrogenase, putative"  
/protein\_id="AAK44926.1"  
/db\_xref="GI:13880223"  
/translation="MSDRHVYVNOVPLENNPASPVLIALLOEGCONCLDEVNEY  
CAISNSCOAKRGELADNRPILHTHDATIGRVDEVEDPVAHELMRLATTHGHAAP  
WADNDRGHVRAAKTSVTVPEGHICISMTYAVVPAALRYNSLAAVETPLTSREY  
DEPLPATTKAGITAGMSMTERKOGSDVRACTOTATPADGYSLTGKMFSAKMD  
IFLVLAQAPDGLSCFLPRVLDPDGRNMFLORLDKLGNHANSSEVEYDGAAMV  
GEEGVPITIEMLVTLRDLCAISATSMRTGLRAVHAQHRKAFGAYLIDPILAMV  
VLADAVEAEATIVAMRAGATNANVGNTEALLIRIGLAAKAYKCRSTHAAB  
ALECIGNGVYEDSGMPLRYEARPLMKIMESGVSALDTIRAAATPACVEVLEDEL  
ARSAQODRLDGHVRLRPQGLDITIGTARRAKIAEDICLALDGSLLVRGHFAVAA  
FLATRLSGVGGAATGMPAGLADLIALERLVK"  
13498. .14436  
/gene="MT0702"  
13498. .14436  
/gene="MT0702"  
/note="similar to GP:3885480; identified by sequence  
similarity; putative"  
/codon\_start=1  
/transl\_table=11  
/product="enoyl-CoA hydratase/isomerase family protein"  
/protein\_id="AAK44927.1"  
/db\_xref="GI:13880224"  
/translation="MTHAIRPVDFDNLKMTYEVTGRARIITNREKNAIADTPL  
ELSLVERADLDGVHVLVSGRGEFCAGDLSAVAESSSTGGGAYGVLDCKT  
OAVNHLPNQWPMIDYOMMSRFVGRFASLYNHAKPTVVVHIGCVAGTDLIHLNDQ  
VIAAADAKIGYPRPVWGVRPAAGLMAHRAKRLFTFGDITGQAAEMGLAVBA  
PEPADLDETERLVARIALPVLNOLIVKALISVALLQGVATSRVSTYFDAAKHT  
PEGHAFVADVEHGFRAVRRRDEPFDYGRQASRV"  
14439. .15161  
/gene="MT0703"  
14439. .15161  
/gene="MT0703"  
/note="Identified by glimmer2; putative"  
/codon\_start=1  
/transl\_table=11  
/product="hypothetical protein"

Query Match 99.2%; Score 206.4; DB 1; Length 19352;  
Best Local Similarity 99.5%; Pred. No. 1,2e-33;  
Matches 207; Conservativity 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 TCAAGGAGAGCGGTACGACCTGGCCCGCGGTCCGTATTAAGCTCAACAAGACTCG 60  
Db 1064 TCAAGGAGAGCGGTACGACCTGGCCCGCGGTCCGTATTAAGCTCAACAAGACTCG 60  
QY 61 GCGTCATGTCGGCGAGCCATCACTGCTGACGCTGACCGAAGAAGACGTGGCCA 120  
Db 1124 GCGTCATGTCGGCGAGCCATCACTGCTGACGCTGACCGAAGAAGACGTGGCCA 1183  
QY 121 CCATCGAATATCTGTCGCGCTTGCACGAGGCTCAACACGATGATCTCCGGCGCGC 180  
Db 1184 CCATCGAATATCTGTCGCGCTTGCACGAGGCTCAACACGATGATCTCCGGCGCGC 1243  
QY 181 TCGAGTCCCGGTGAACACGACGACAT 208  
Db 1244 TCGAGTCCCGGTGAACACGACGACAT 1271

RESULT 14  
LOCUS MT01376 19770 bp DNA linear BCT 03-ANG-2001  
DEFINITION Mycobacterium tuberculosis H37Rv complete genome; segment 32/162.  
ACCESSION 295972 AL123456  
VERSION 295972.1 GI:3261790  
KEYWORDS  
SOURCE Mycobacterium tuberculosis H37Rv.  
ORGANISM Mycobacterium tuberculosis H37Rv.  
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
Mycobacterium; Mycobacteriaceae; Mycobacteriales; Corynebacterineae; Mycobacteriaceae;  
Mycobacterium; Mycobacterium tuberculosis complex.

REFERENCE 1 (bases 1 to 19770)  
 AUTHORS Cole, S.T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C., Harris, D., Gordon, S.V., Eigmeier, K., Gas, S., Barry III, C.E., Tekala, P., Badcock, K., Basham, D., Brown, D., Chillingworth, T., Connor, R., Davies, R., Devlin, K., Fellwell, T., Gentles, S., Hamlin, N., Holroyd, S., Hornsby, T., Jagels, K., Krogh, A., McLean, J., Mole, S., Murphy, L., Oliver, S., Osborne, J., Quail, M.A., Rajadaram, M.A., Rogers, J., Ruter, S., Seeger, K., Skelton, S., Squares, S., Squires, R., Sulston, J.E., Taylor, K., Whitehead, S., and Barrall, B.G.  
 TITLE Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence  
 JOURNAL Nature 393 (6685), 537-544 (1998)  
 MEDLINE 98295987  
 PUBMED 9634230  
 REFERENCE 2 (bases 1 to 19770)  
 AUTHORS Parkhill, J.  
 TITLE Direct Submission  
 JOURNAL Submitted (11-JUN-1998) Submitted on behalf of the Mycobacterium tuberculosis sequencing and mapping teams, Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA Unite de Genetique Moleculaire Bacterienne, Institut Pasteur, 28 rue du Docteur Roux, 75724 Paris Cedex 15, France E-mail: parkhill@sanger.ac.uk  
 COMMENT On Jun 27, 1998 this sequence version replaced gi:2143285.  
 Notes:  
 Details of *M. tuberculosis* sequencing at the Sanger Centre are available on the World Wide Web.  
 (URL, [http://www.sanger.ac.uk/Projects/M\\_tuberculosis/](http://www.sanger.ac.uk/Projects/M_tuberculosis/)) CDS have been renumbered from the original cosmid submissions but the old gene designations are in brackets after the new gene numbers.  
 Gene prediction was based on a Hidden Markov Model of 78 genes implemented in Reparse (Krogh) supplemented with visual inspection of positional base preference in codons, especially where there is an increase in the observed/expected third position G + C.  
 CAUTION: In some cases we may not have predicted the correct initiation codon. Where possible we choose an initiation codon (atg, gtg, or ttg) which is preceded by an upstream ribosome binding site sequence (optimally 5-13bp before the initiation codon). If this cannot be identified we choose the most upstream initiation codon.  
 Location/Qualifiers  
 source 1..19770  
 /organism="Mycobacterium tuberculosis H37Rv"  
 /strain="H37Rv"  
 /db\_xref="taxon:83332"  
 /clone="1376"  
 57..61  
 /note="agga, possible rbs upstream of Rv0654"  
 68..1573  
 /gene="Rv0654"  
 68..1573  
 /note="Rv0654" (MTC1376.22), len: 501, unknown, FASTA score: 0.53153 LIGNOSTILBENE-ALPHA, BETA-DIOXYGENASE (485 aa) opt:280 z-score: 330.1 E(): 2.3e-11, (28.5% identity in 523 aa overlap). Also similar to *M. tuberculosis* protein MTCY21C12\_07c (29.5% identity in 522 aa overlap).  
 /codon\_start=1  
 /transl\_table=1  
 /product="hypothetical protein Rv0654"  
 /protein\_id="CAB09380.1"  
 /db\_xref="GI:2143307"  
 /db\_xref="SPTREMBL:006785"  
 /translation="MTTAAQASQNPYLGGFLAPSTEVATDLPVYTGRIPEHLDGRYLRNPNVAEVDPAATYTHPTGDMAGVAGVARDKAKWYNNKVRTPAVCAALGEISARPRRTGIIIGGFNTNVLTHAGRTALVEAGVAVNELTDELDTGVCPCDFGTLHGAYTAVVYIDLPVTFDMQVAVSFARGHRYOVSIGTGHARRTVDIIVAGSPMHSFSLTDNRLDYAMPSTPARVGVMPREGNEDVWMPDIEPCVYHPLNAYSECRRNGAEVLYVHYRYSRMPDRDRGGDSRPSLDKWTINLATGAVTAACRRDRQERPRINETIVGSPHRYFAYTIGEGFLVAGCAALSTPIKDCDCTGSSYVASLDPDLIGEMVFPVPSRAEDDGIILMGCHMRDRDGGQLLLDAOTLESIAIVHLPORVPMGFHGNMAPPY"  
 1585..2664  
 /gene="Rv0658c" (MTC1376.18), len: 238, unknown,

CDS  
 /gene="Rv0655"  
 1585..2664  
 /gene="Rv0655"  
 /note="Rv0655" (MTC1376.21), len: 359, abc transporter, FASTA score: TRBF\_ECOLI P45393 hypothetical abc transporter atp-binding (269 aa) opt: 644 z-score: 721.8 E(): 3.4e-33 (38.5% identity in 244 aa overlap); contains PS00017 ATP/GTP-binding site motif A, PS00211 ABC transporters family signature, highly similar to *M. leprae* MKL\_MYCCE P30769 possible ribonucleotide transport atp- in 335 aa overlap). Also similar to many other *M. tuberculosis* ABC transporters eg. MTCY253.24 (33.6% identity in 241 aa overlap).  
 /codon\_start=1  
 /transl\_table=1  
 /product="hypothetical protein Rv0655"  
 /protein\_id="CAB09379.1"  
 /db\_xref="GI:2143306"  
 /db\_xref="SWISS-PROT:006784"  
 /translation="MRYSDSYHTGRMQPRASTEGFPMGVSIENVGLTKSGSRIME DVTLP:PGSEVSLGPGTGKSVELKSLGLRPERGSITIGDITIGSAEIKETI RTLFGVLPRDGLGSMNLVNTAPFPIREHKKESIRIYVEKALVGLGDEKFI PGEISGHRKAGARALVLDPOIILDEPDGSLDPVRYVLSQLMINDNAQIDATIL IYTHNINARVPPDMGMLEFKHLVMTGPEVRLTSDPVRQFNLNGRIGPIGMSEEDKATMAEQALLDLGHHAGVEIEGVPPQISATPMPERKAVARQARVREMLHTL PKKAVALILDOI:EGTHKVAHEIQ"  
 1762..1785  
 /gene="Rv0655"  
 /note="PS00017 ATP/GTP-binding site motif A"  
 2074..2118  
 /gene="Rv0655"  
 /note="PS00211 ABC transporters family signature"  
 complement(3052..3435)  
 /gene="Rv0656c"  
 complement(3052..3435)  
 /gene="Rv0656c"  
 /note="Rv0656c" (MTC1376.20, unknown), len: 127 aa  
 /codon\_start=1  
 /transl\_table=1  
 /product="hypothetical protein Rv0656c"  
 /protein\_id="CAB09378.1"  
 /db\_xref="GI:2143305"  
 /db\_xref="SPTREMBL:006783"  
 /translation="MAAATTGTHRGLELRAQRAVSGCEPORAEPFCRARNADEFDQ MSRRGDIYDPVPEPKSVWRMIDSAQRIARARAGVAGLSVYDILLICDTAAAGLVYLIH DDADYELAEHLDPIDIRVRYSAD"  
 complement(3530..3685)  
 /gene="Rv0657c"  
 complement(3530..3685)  
 /gene="Rv0657c"  
 /note="Rv0657c" (MTC1376.19), unknown, len: 51 aa; similar to several other *M. tuberculosis* hypothetical proteins eg. YW08\_MYCTU Q10848 hypothetical 8.9 kd protein c739\_08c (80 aa), FASTA scores: opt: 107 z-score: 182.3 E(): 0.0039, 45.8% identity in 48 aa overlap. Also similar to MTCY48\_5 aa), 41.0% identity in 39 aa overlap.  
 /codon\_start=1  
 /transl\_table=1  
 /product="hypothetical protein Rv0657c"  
 /protein\_id="CAB09377.1"  
 /db\_xref="GI:2143304"  
 /db\_xref="SPTREMBL:006782"  
 /translation="MSYQIDLDLDELADVMRIAAVHTKKEAVNLAMDYERPRRIE ALARSR"  
 complement(3691..3696)  
 /note="possible RBS upstream of Rv0657c"  
 complement(3761..4477)  
 /gene="Rv0658c"  
 complement(3761..4477)  
 /gene="Rv0658c"  
 /note="Rv0658c" (MTC1376.18), len: 238, unknown,

gene

probablemembrane protein, similar to YPRB\_ECOLI\_P33774  
hypothetical24.3 kd protein (urf 1) (217 aa), fasta  
scores: opt: 174 z-score: 215.8 E(): 5.3e-05, (25.6%  
identity in 223 aa overlap). Also similar to MTCY359.10  
(28.7% identity in 178 aa overlap)"

/codon\_start=1

/transl\_table=11

/product="hypothetical protein Rv0658c"

/protein\_id="CA09376.1"

/db\_xref="GI:2143303"

/db\_xref="SPTREMBL:O06781"

/translation="MEGRADYVAPSHRWGLGAPLVVELVFLVASTSLAVVTGHPV  
SAGVLALALAPVVAAGLAILITRLRGLRSMNGLGLMPPGGALVT  
IPASLYTAVIGPEANSVAVRIEGGRASAPALVPLVYVVAAPICEITTYRGLMG  
AVDRMGKNAALVYTVTFALAHLEPARAPLVVAIPALARFISGGLASIVTHOV  
TNLPGLVILLGTTGASISP"

complement(4480..4483)

/note="possible RBS upstream of Rv0658c"

complement(4753..5061)

/gene="Rv0659c"

complement(4753..5061)

/note="Rv0659c, (MTCI376.17), len: 102; unknown, similar  
to YW28\_MCTU\_Q10867 hypothetical 12.3 kd protein cy3.28  
(114 aa), fasta scores: opt: 144 z-score: 213.2 E():  
7.3e-05, 30.8% identity in 107 aa overlap. Also similar to  
MTCY09F9.22 (32.7% identity in 101 aa overlap)"

/codon\_start=1

/transl\_table=11

/product="hypothetical protein Rv0659c"

/protein\_id="CA09387.1"

/db\_xref="GI:2143302"

/db\_xref="SPTREMBL:O06780"

/translation="MRGELMFAATPGDRPLVLTFRPVDRIQAVVVAATPFRG  
LVSELELTAVENRVPDVCVNFNDNIHTLPRFAFRIRITRLSPALHENCOTLRASGTC  
"

complement(5048..5293)

/gene="Rv0660c"

complement(5048..5293)

/gene="Rv0660c"

/note="Rv0660c, (MTCI376.16), len: 81; some similarity to  
IAF016485.130 Halobacterium sp: NRC-1 plasm (100 aa),  
32.4% identity in 74 aa overlap"

/codon\_start=1

#### Query Match

Best Local Similarity 99.2%; Score 206.4; DB 1; Length 19770;

Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
OY 1 TCAAGGAGAGCGCTACGACCTGCGCCGCGTGGTCCGTATTAAGTCAACAAGAGCTCG 60
    |||||||
Db 10758 TCAAGGAGAGCGCTACGACCTGCGCCGCGTGGTCCGTATTAAGTCAACAAGAGCTCG 10817
OY 61 GGCTGATGTGCGGAGCCCATCAGCTGCTGACCGCTGACCAAGAGACGTCGTGCCA 120
    |||||||
Db 10818 GGCTGATGTGCGGAGCCCATCAGCTGCTGACCGCTGACCAAGAGACGTCGTGCCA 10877
OY 121 CCATCGAATATCTGCTCCGCTTGACAGAGGTCACACCAAGATGATCTCCGGCGCGC 180
    |||||||
Db 10878 CCATCGAATATCTGCTCCGCTTGACAGAGGTCACACCAAGATGATCTCCGGCGCGC 10937
OY 181 TCGAGTGCCGGTGAAGACCGAGACAT 208
    |||||||
Db 10938 TCGAGTGCCGGTGAAGACCGAGACAT 10965
```

RESULT 15  
AR067447

LOCUS AR067447 3447 bp DNA linear PAT 29-SEP-1999

DEFINITION Sequence 57 from patent US 5851763.

ACCESSION AR067447  
VERSION AR067447.1 GI:5998669

#### SOURCE

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 3447)

AUTHORS Hymn B., Cole, S., Young, D., Zhang, Y., Honore, N., Telenti, A. and

Bohmer, T.

TITLE Rapid detection of antibiotic resistance in mycobacterium

JOURNAL Patent: US 5851763-A 57 22-DEC-1998;

FEATURES location/Qualifiers

source 1..3447

BASE COUNT 687 a 965 c 1139 g 656 t

ORIGIN

Query Match 70.8%; Score 147.2; DB 6; Length 3447;

Best Local Similarity 81.7%; Pred. No. 3e-21; Mismatches 38; Indels 0; Gaps 0;

Matches 170; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

```
OY 1 TCAAGGAGAGCGCTACGACCTGCGCCGCGTGGTCCGTATTAAGTCAACAAGAGCTCG 60
    |||||||
Db 809 TCAAGGAGAGCGCTACGACCTGCGCCGCGTGGTCCGTATTAAGTCAACAAGAGCTCG 868
OY 61 GGCTGATGTGCGGAGCCCATCAGCTGCTGACCGCTGACCAAGAGACGTCGTGCCA 120
    |||||||
Db 869 GGCTGATGTGCGGAGCCCATCAGCTGCTGACCGCTGACCAAGAGACGTCGTGCCA 928
OY 121 CCATCGAATATCTGCTCCGCTTGACAGAGGTCACACCAAGATGATCTCCGGCGCGC 180
    |||||||
Db 929 CCATCGAATATCTGCTCCGCTTGACAGAGGTCACACCAAGATGATCTCCGGCGCGC 988
OY 181 TCGAGTGCCGGTGAAGACCGAGACAT 208
    |||||||
Db 989 TCGAGTGCCGGTGAAGACCGAGACAT 1016
```

Search completed: November 13, 2002, 01:27:59  
Job time : 660.723 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 12, 2002, 16:59:42 ; Search time 1141.13 Seconds

(without alignments)  
3037.202 Million cell updates/sec

Title: US-09-697-123b-21

Sequence: 1 tcaagagagaagcgcacgat.....ccggtgagatgtgacacacat 214

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*  
1: em\_estba:\*  
2: em\_estlin:\*  
3: em\_estlin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hlc:\*  
9: gb\_estcl:\*  
10: gb\_estcl:\*  
11: gb\_hlc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vit:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rtd:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	137.6	64.3	1282	9 A1770311	A1770311.42 Mycoba
C 2	47.4	22.1	925	17 CNS0091P	AL053013 Drosophi1
C 3	46	21.5	417	14 BM85094	BM85094 mgap008xj
C 4	44.6	20.8	569	17 BH897054	BH897054 3526_1_6
5	44	20.6	585	13 BM489075	BM489075 pgnm2.pk0
6	44	20.6	635	13 B1065677	B1065677 pgtfn.pk0

7	44	20.6	666	13 BM486272	BM486272 pgn1c.pk0
8	44	20.6	669	13 BM485879	BM485879 pgn1c.pk0
9	44	20.6	672	13 B1716865	B1716865 103101380
10	44	20.6	682	13 BM485533	BM485533 pgn1n.pk0
11	44	20.6	783	12 BF864011	BF864011 963048D06
12	43.8	20.5	568	13 B1995346	B1995346 103102680
13	43.4	20.3	710	12 BG321153	BG321153 zmo4_05f0
14	43.2	20.2	1856	11 AY109418	AY109418 zea mays
15	43	20.1	558	12 BE754521	BE754521 zea mays
16	42.8	20.0	275	10 R63542	R63542 RABEST103T
17	42.6	19.9	455	10 AV635603	AV635603 AV635603
C 18	42.6	19.9	523	14 A1856228	A1856228 sb39e01.x
19	42.4	19.8	487	14 B0824742	B0824742 tetradon
20	42.4	19.8	923	17 CNS03VB3	CNS03VB3
C 21	42.2	19.7	553	9 A178937	A178937 zpj9c04.s
22	42	19.6	373	12 BG159643	BG159643 OVE2_7_C05
23	42	19.6	370	10 AWE64128	AWE64128 LG1_282_E
24	42	19.6	477	14 W17149	W17149 zbl3b04.r1
25	42	19.6	593	10 BE358685	BE358685 DC1_31_D0
C 26	42	19.6	710	13 BG933256	BG933256 W81_92_B1
27	42	19.6	2598	11 AY103647	AY103647 zea mays
28	41.6	19.4	373	10 BE442717	BE442717 WHE1105_C
29	41.6	19.4	396	10 BE429399	BE429399 MTD017.F0
30	41.6	19.4	515	12 BF177122	BF177122 EM1_2_A10
31	41.6	19.4	587	14 B0743581	B0743581 WHE4105_F
32	41.6	19.4	636	14 B0294664	B0294664 WHE2853_A
33	41.4	19.3	234	14 R86588	R86588 RABEST163T
34	41.4	19.3	235	14 R86566	R86566 RABEST140T
35	41.4	19.3	265	14 E27471	E27471 HSPD15421.H
36	41.4	19.3	318	14 F00050	F00050 HSB09E061.S
37	41.4	19.3	331	14 R86510	R86510 RABEST202T
38	41.4	19.3	369	14 F36189	F36189 HSPD33571.H
39	41.4	19.3	376	14 F21475	F21475 HSPD06020.H
40	41.4	19.3	389	14 F25575	F25575 HSPD12678.H
41	41.4	19.3	392	9 AA180106	AA180106 zpj94907.f
42	41.4	19.3	397	9 AA112975	AA112975 zn59c09.f
43	41.4	19.3	475	12 BF221644	BF221644 7059h10.x
C 44	41.4	19.3	502	9 AA211560	AA211560 zn55f10.f
45	41.4	19.3	502	14 BM698942	BM698942 UI-E-DX1-

## ALIGNMENTS

RESULT 1  
A1770311/c  
LOCUS A1770311 1282 bp mRNA linear EST 24-JAN-2000  
DEFINITION 42 Mycobacterium anaerobic stationary phase library Mycobacterium  
smegmatis cDNA, mRNA sequence.  
ACCESSION A1770311  
VERSION A1770311.1 GI:6742680  
KEYWORDS EST.  
SOURCE Mycobacterium smegmatis.  
ORGANISM Mycobacterium smegmatis.  
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
Corynebacterineae; Mycobacteriaceae; Mycobacterium.

REFERENCE  
1 (bases 1 to 1282)  
Murusu-Oel,B., Tay,A. and Dick,T.  
Upregulation of stress response genes and ABC transporters in  
anaerobic stationary-phase Mycobacterium smegmatis  
Mol. Gen. Genet. 262 (4-5), 677-682 (1999)  
JOURNAL MEDLINE  
20092472  
COMMENT Contact: Murusu-Oel, B.  
Mycobacterium Laboratory  
Institute of Molecular and Cell Biology  
30 Medical Drive, Singapore 117609, Republic of Singapore  
Tel: 65 874 3011  
Fax: 65 779 1117  
Email: mcbomelmbc.nus.edu.sg  
Insert length: 1282 Std Error: 0.00  
Seq primer: T3 Forward: T7 Backward.  
Location/Qualifiers  
1..1282

FEATURES  
source









RESULT 8  
BM485879  
LOCUS  
DEFINITION  
669 bp mRNA linear EST 07-FEB-2002  
pmlc.pk002.ag Primary Chicken Breast Muscle, Leg Muscle, and  
Epiphyseal Growth Plate CDNA Library (pmlc) Gallus gallus CDNA  
clone pmlc.pk002.ag 5' similar to sp|P025881|TPCS\_CHICK\_TROPONIN\_C,  
SKELETAL MUSCLE p1l7PCHCS tropoin C, skeletal muscle - chicken,  
mRNA sequence.  
BM485879  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
Phasianinae; Gallus.  
EST.  
BM485879.1 GI:18606809  
SOURCE  
ORGANISM  
chicken.  
Gallus gallus  
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
Phasianinae; Gallus.  
REFERENCE  
AUTHORS  
TITLE  
1 (bases 1 to 669)  
Coburn,L.A. and Monsonego-Ornan,E.  
ESTs from Primary Chicken Breast Muscle, Leg Muscle, and Epiphyseal  
Growth Plate CDNA Library, USDA/IRAFs Animal Genome Project  
Unpublished (2002)  
JOURNAL  
COMMENT  
Contact: Larry A. Coburn  
University of Delaware  
Towmsend Hall, Newark, DE 19717, USA  
Tel: 302-831-1335  
Fax: 302-831-2822  
Email: coburn@udel.edu, www.chickest.udel.edu.  
FEATURES  
source  
1. 669  
Location/Qualifiers  
/organism="Gallus gallus"  
/strain="Commercial broiler and Ottawa Res. Centre  
Strains 90 & 21"  
/db\_xref="taxon:9031"  
/clone="pmlc.pk002.ag"  
/clone.lib="Primary Chicken Breast Muscle, Leg Muscle, and  
Epiphyseal Growth Plate CDNA Library (pmlc)"  
/sex="Male and Female"  
/tissue="type="Breast muscle, leg muscle and epiphyseal  
growth plate"  
/dev\_stage="Breast,leg:Embryo(d19);post-hatch(1d,1,3,5,7,9  
/lab\_host="E. coli EMH10B"  
/note="Vector: PCMVSPORT6; Library made from equivalent  
pools of total RNA isolated from each tissue (embryonic  
muscle 33.3%; juvenile muscle 33.3%; and epiphyseal growth  
plate 33.3% of the final RNA pool). Single pass sequencing  
from 5'-end"  
BASE COUNT 152 a 172 c 230 g 112 t 3 others  
ORIGIN  
Query Match 20.6%; Score 44; DB 13; Length 669;  
Best Local Similarity 53.5%; Pred. No. 6.2;  
Matches 92; Conservative 0; Mismatches 80; Indels 0; Gaps 0;  
OY 41 CAAGGTGAAACAAGACTGGGCTGGGGCGACCAATCCGGCTCAGGTACCAACACAC 100  
DB 169 CAGGAGATGAGATGCTGGGCGACCAAGCCACCAAGAGAGAGCTGATGCATCATCA 228  
OY 101 CCTCAGCAGAGAGAGCTGCGCCACATCGACTGCTGGGCGCTGCACAGAGGCCA 160  
DB 229 GAGAGTGGAGAGATGGCAGCGGCGACCATCGACTTCGAGAGAGTTCTTGATGATG 288  
OY 161 GACCACGATGACCGCCCGCGGCGCTGAGAGTCCGCTGATGTGGACGAC 212  
DB 289 GCCCAGATGAAGAAGAGGCGCCCAAGGCAAGTCTGAGAGAGAGCTGCCAAC 340  
RESULT 9  
B1716865  
LOCUS  
DEFINITION  
672 bp mRNA linear EST 19-SEP-2001  
1031013B08.y1 C. reinhardtii CC-1690, Stress II (normalized),  
Lambda Zap II Chlamydomonas reinhardtii CDNA, mRNA sequence.  
ACCESSION  
B1716865

VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
DCMB Box 91000  
Duke University  
Durham, NC 27708-1003  
Tel: 919 613 8159  
Fax: 919 613 8177  
Email: chausser@duke.edu.  
FEATURES  
source  
1. 672  
Location/Qualifiers  
/organism="Chlamydomonas reinhardtii"  
/strain="CC-1690 wild type mt+ 21gr"  
/db\_xref="taxon:3055"  
/clone.lib="C. reinhardtii CC-1690, Stress II (normalized  
), Lambda Zap II"  
/note="Vector: pluescript II SK-; Site\_1: EcoRI; Site\_2:  
XhoI; Stress condition II library, constructed by John  
Davies and Jeffrey McDermott, combines cDNAs from CC-1690  
cells grown to mid-log phase in TAP (NH4+ - containing)  
and shifted to TAP - NO3- (24hrs); h2 production  
conditions (0, 12hr, 24hr) see Wells et al., (2000) Plant  
Phys. 122: 127-135; TAP + H2O2 (1, 12, 24 hr); TAP +  
sorbitol (1, 2, 6, 24 hr); TAP + Cd (1, 2, 6, 24 hr).  
PolyA mRNA was purified from each sample, pooled and CDNA  
synthesized. The cDNA was directionally cloned into lambda  
Zap II (Stratagene) in the EcoRI (5') and XhoI (3')  
sites. pluescript II SK- plasmids were excised from the  
lambda Zap clones by superinfection with Exsist  
(Stratagene) phage. The library was normalized using  
method 4 described in Bonaldo et al., (1996) Genome  
Research 6: 791-806."  
BASE COUNT 131 a 207 c 224 g 110 t  
ORIGIN  
Query Match 20.6%; Score 44; DB 13; Length 672;  
Best Local Similarity 52.8%; Pred. No. 6.2;  
Matches 95; Conservative 0; Mismatches 85; Indels 0; Gaps 0;  
OY 35 TCGGTACAAGGTGAACAAGCTGGGCGCGCGCCACCAATCCGGCTCAGGTGACAC 94  
DB 353 TCCGAGAAATGACCAAAATGTGGCGAGGCGGCGCTTTCAGTGTGGCGCCAAA 412  
OY 95 CACCACTCTCACCGAGAAAGCTGTCGCCACATCAGTACGTGTCGCTGCACGA 154  
DB 413 GCAGGCGGTGGCAATGGCGCGCGCGCTGTGAAGAACCCCAAGCTCTTCTGGAGCA 472  
OY 155 GGGCCAGACCAAGATGACCGCCCGCGGCGCTGAGGTGGGTGATGTGACGACAT 214  
DB 473 GGCACATGTCGGCGTGAAGCGCCCGCAGAGCGAGCGGTGTGACGCGGCGTGCACCG 532  
RESULT 10  
BM485533  
LOCUS  
DEFINITION  
682 bp mRNA linear EST 07-FEB-2002  
pmln.pk001.06 Normalized Chicken Breast Muscle, Leg Muscle, and  
Epiphyseal Growth Plate CDNA Library (pmln) Gallus gallus CDNA  
clone pmln.pk001.06 5' similar to sp|P025881|TPCS\_CHICK\_TROPONIN\_C,  
SKELETAL MUSCLE p1l7PCHCS tropoin C, skeletal muscle - chicken,  
mRNA sequence.  
ACCESSION  
BM485533  
VERSION  
BM485533.1 GI:18606186

KEYWORDS	EST.
SOURCE	Chicken
ORGANISM	Gallus gallus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus. 1 (bases 1 to 682)
AUTHORS	Cogburn,L.A. and Monsonego-Ornan,E.
TITLE	ESTs from Normalized Chicken Breast Muscle, Leg Muscle, and Epiphyseal Growth Plate cDNA library, USDA/IFAFS Animal Genome Project
JOURNAL	Unpublished (2002)
COMMENT	Contact: Larry A. Cogburn University of Delaware Townsend Hall, Newark, DE 19717, USA Tel: 302-831-1335 Fax: 302-831-2822 Email: cogburn@udel.edu, www.chickest.udel.edu.
FEATURES	Location/Qualifiers
source	1..682
/organism="Gallus gallus"	
/strain="Commercial broiler and Ottawa Res. Centre	
Strains 90.6.21"	
/db_xref="taxon:9031"	
/clone="pgmln.pk001.o6"	
/clone_1lb="Normalized Chicken Breast Muscle, Leg Muscle,	
and Epiphyseal Growth Plate cDNA library (pgmln)"	
/sex="Male and Female"	
/tissue_type="Breast muscle, leg muscle and epiphyseal	
growth plate"	
/dev_stage="Breast,leg:Embryo(d19);post-hatch(1d,1,3,5,7,9,	
11 weeks);growth plate(1d,7d,14d post-hatch)"	
,lab_host="E. coli EMDH10B"	
/note="Vector: pCMVSPORT6; Library made from equivalent	
pools of total RNA isolated from each tissue (embryonic	
muscle 33.3%; juvenile muscle 33.3%; and epiphyseal growth	
plate 33.3% of the final RNA pool). Single pass sequencing	
from 5'-end"	
BASE COUNT	151 a 179 c 230 g 120 t 2 others
ORIGIN	
Query Match	20.6%; Score 44; DB 13; Length 682;
Best Local Similarity	53.5%; Pred. No. 6.3;
Matches	92: Conservative 0; Mismatches 80; Indels 0; Gaps 0;
Oy	41 CAGGTGAACAAGAAGCTGGCGCTGGCGGCACCAATCCGGCTCAGTGCACCACCAC 100
Db	162 CACGCTGATGAGGATGCTGGGCCAACAMCCCACCAAAGAGCATGTCATCATCGA 221
Oy	101 CCTCACCGAGAAAGCGTCGCGCCACCATCAGTAGTAACCTGGTGCCCTGCACAGAGGCCA 160
Db	222 GGAGGTGACACAGATGTCACAGCGCACCATTGCATCTTCAGAGGATTTCTGTGATATATGT 281
Oy	161 GACCACGATGACCCCGCGCGCGCGGTGAGAGTGGCGGTGATGTGCAGCAC 212
Db	282 GCGCGCATGGAAGAGAGCGCCAAAGGCAAGTCGTGAGAGAGAGCTGGCCAC 333
RESULT 11	
Locus	BF864011
DEFINITION	BF864011 783 bp mRNA linear EST 19-JAN-2001
ACCESSION	963048DD06.y1 C. reinhardtii CC-1690, Stress condition I, normalized
VERSION	BF864011
KEYWORDS	, Lambda zap II Chlamydomonas reinhardtii cDNA, mRNA sequence.
SOURCE	BF864011.1 GI:12254155
ORGANISM	EST.
Chlamydomonas reinhardtii.	
Chlamydomonas reinhardtii	
Eukaryote; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;	
Chlamydomonadaceae; Chlamydomonas.	
1 (bases 1 to 783)	
Grossman,A., Davies,J., Federspiel,N., Harris,E., Hauser,C.,	
Lefebvre,P., McDermott,J.P., Shrago,J., Silflow,C. and Stern,D.	

TITLE	Analyses of the Chlamydomonas reinhardtii Genome: A Model, Unicellular System for Analyzing Gene Function and Regulation in Vascular Plants, project phase 3		
JOURNAL	Unpublished (2000)		
COMMENT	Contact: Charles Hauser DCMB Box 91000 Durham, NC 27708-1000 Tel: 919 613 8159 Fax: 919 613 8177 Email: chauser@duke.edu.		
FEATURES	Location/Qualifiers		
source	1. 783 /organism="Chlamydomonas reinhardtii" /strain="CC-1690 wild type mt+ 219r" /db_xref="taxon:3055" /clone_11b="C. reinhardtii CC-1690, Stress condition 1, normalized, Lambda zap II" /note="Vector: pBluescript II SK-; Site.1: EcoRI; Site.2: XhoI. This library, constructed by John Davies and Jeffrey McDermott, combines cDNAs from CC-1690 cells grown to mid-log phase in TAP-N (30 min, 1hr, 4hr), TAP-S (30 min, 1hr, 4hr), TAP-P (4hr, 12hr, 24hr), NO3 to NH4 (30min, 1hr, 4hr) and NH4 to NO3 (30min, 1hr, 4hr). PolyA mRNA was purified from each sample, pooled and cDNA synthesized. The cDNA was directionally cloned into lambda zap II (Stratagene) in the EcoRI (5') and XhoI (3') sites. pBluescript II SK- plasmids were excised from the lambda zap clones by superinfection with Exassist (Stratagene) phase. The library was normalized using method 4 described in Bonaldi et al (1996) Genome Research 6: 791-806."		
BASE COUNT	167 a 226 c 256 g 132 t 2 others		
ORIGIN			
Query Match	20.6%; Score 44; DB 12; Length 783;		
Beat Local Similarity	52.8%; Pred. NO. 6.4;		
Matches	95: Conservative 0; Mismatches 85; Indels 0; Gaps 0;		
QY	35 TCGGTACAAAGTGAACAGAGCTGGCGCGGCACCAATCCGGCTCAGTGACAC 94		
DB	386 TTCGGAGAGTACACACCAAAATGTGGCGGCGGCGCATTCAGTGAAGTGGCGGCCAAA 445		
QY	95 CACCAACCCCTACCGAGGAAGAGCTGTGGCCACCTCAGTACTCTGTGGCTTCAGCA 154		
DB	446 GCAGGCGGTGGCCATTGGCGCGCGTGTGAAGAACCCCAAGCTCTTGTCTTGAGAGA 505		
QY	155 GGGCCAGACCAAGATGACCGCCCGCGCGCTCAGAGGCGCGGTGATGTGGAGACAT 214		
DB	506 GGCAACGTGGGCGCTTAGAACGCCCGCAGCAGCGGTGTGTGACGGCGCGCTGAGCCGAT 565		
RESULT 12			
LOCUS	B1995346 568 bp mRNA linear EST 25-OCT-2001		
DEFINITION	1031026E07.Y2 C. reinhardtii CC-1690, Stress II (normalized), Lambda zap II Chlamydomonas reinhardtii cDNA, mRNA sequence.		
ACCESSION	B1995346		
VERSION	B1995346.1 GI:16430136		
KEYWORDS	EST.		
SOURCE	Chlamydomonas reinhardtii.		
ORGANISM	Chlamydomonas reinhardtii.		
REFERENCE	Eukaryote: Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales; Chlamydomonadaceae; Chlamydomonas.		
AUTHORS	1 (bases 1 to 568) Grossman,A., Chang,C.-W., Davies,J., Harris,E., Hauser,C., Lefebvre,		
TITLE	'P., McDermott,J.P., Shrago,J., Silflow,C. and Stern,D.		
JOURNAL	Analyses of the Chlamydomonas reinhardtii Genome: A Model,		
COMMENT	Unicellular System for Analyzing Gene Function and Regulation in Vascular Plants. Project: 1031		
COMMENT	Unpublished (2001)		
COMMENT	Contact: Charles Hauser		
COMMENT	DCMB Box 91000		
COMMENT	Duke University		

Durham, NC 27708-1000  
 Tel: 919 613 8159  
 Fax: 919 613 8177  
 Email: chauser@duke.edu.  
 Location/Qualifiers  
 1.568

FEATURES  
 source  
 /organism="Chlamydomonas reinhardtii"  
 /strain="CC-1690 wild type mt+ 21gr"  
 /db\_xref="taxon:3055"  
 /clone\_1lb="C. reinhardtii CC-1690, Stress II (normalized)  
 ), Lambda Zap II"  
 /note="vector: pbluescript II SK-; Site\_1: EcoRI; Site\_2:  
 XhoI; Stress condition II library, constructed by John  
 Davies and Jeffrey McDermott, combines cDNAs from CC-1690  
 cells grown to mid-log phase in TAP (NH4+ - containing)  
 and shifted to TAP - NO3- (24hrs); H2 production  
 conditions (0, 12hr, 24hr) see Mellis et al., (2000) Plant  
 Phys. 122: 127-135; TAP + H2O2 (1, 12, 24 hr); TAP +  
 sorbitol (1, 2, 6, 24 hr); TAP + Cd (1, 2, 6, 24 hr).  
 polyA mRNA was purified from each sample, pooled and cDNA  
 synthesized. The cDNA was directionally cloned into lambda  
 Zap II (Stratagene) in the EcoRI (5') and XhoI (3')  
 sites. pbluescript II SK- plasmids were excised from the  
 lambda Zap clones by superinfection with ExAssist  
 (Stratagene) phage. The library was normalized using  
 method 4 described in Bonaldo et al., (1996) Genome  
 Research 6: 791-806."  
 BASE COUNT 80 a 206 c 185 g 97 t  
 ORIGIN

Query Match 20.5%; Score 43.8; DB 13; Length 568;  
 Best Local Similarity 54.0%; Pred. No. 6.7;  
 Matches 115: Conservative 0; Mismatches 92; Indels 6; Gaps 1;

OY 2 CAAGGAGAGCCCTGATCTGGCCCGGCTGCTGACAGTGAACAAGAGCTGGG 61  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 DB 251 CAAGGAGAGCCGCTTCTGAGTGCTGACAGGCGCTTGAACAGGATGAGAGCTGGG 310  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 OY 62 CTTGGGGCGGACCCATCCGCTCAGTACACCAACCTTCACGAGAAGAGCTGTG 121  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 DB 311 CTTG-----CCGTACGAGGAGACCCGCGGCTCTCCCGCCCTGCTCACACTG 364  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 OY 122 CGCCACATCGATGATCTGCTGCTGACAGGCGGCGGCTGACGATGACGCGCGCGG 181  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 DB 365 GCGCGCGCGCGCGGACCCCGGCGGCTTCCCGGCGGCGGCGGCGGCGGCGGCGGCGG 424  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 OY 182 CGGCTCAGAGTGGCGGCTGATGTGACGACAT 214  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 DB 425 CTGCGCGGACCAAGCCCGCTGCGCTTCCGCAAT 457  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 13  
 BG321153 710 bp mRNA linear EST 27-FEB-2001  
 LOCUS  
 DEFINITION Zm04\_05f06\_R Zm04\_AAF-C\_ECORC\_cold\_stressed\_maize\_seedlings Zea mays  
 BG321153  
 ACCESSION  
 VERSION BG321153.1 GI:13150831  
 KEYWORDS EST.  
 SOURCE Zea mays.  
 ORGANISM Zea mays.  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC  
 clade; Panicoidae; Andropogoneae; Zea.  
 1 (bases 1 to 710)  
 SINGH, J.A., MAKUL, K., COURTOUX, P., DE MOORS, A., HARRIS, L.J., HATTORI  
 J.I., OUELLET, F., ROBERT, L.S., SPROTT, D., and TINKER, N.A.  
 TITILE Expressed Sequence Tags from Cold-Stressed Maize Seedlings  
 JOURNAL Unpublished (2001)  
 COMMENT Contact: Singh, J.A.  
 Eastern Cereal and Oilseed Research Centre  
 Agriculture and Agri-Food Canada  
 KM Mealy Bldg., Central Experimental Farm, Ottawa, Ontario, K1A

OC6, Canada  
 Tel: (613) 759-1662  
 Fax: (613) 759-1701  
 Email: singh@cem.agr.ca.  
 Location/Qualifiers  
 1.710

FEATURES  
 source  
 /organism="Zea mays"  
 /cultivar="C0328"  
 /db\_xref="taxon:4577"  
 /clone="Zm04\_05f06"  
 /clone\_1lb="Zm04\_AAF-C\_ECORC\_cold\_stressed\_maize\_seedlings"  
 /tissue\_type="leaf, crown"  
 /note="vector: Bluescript SK-/XhoI-EcoRI; Site\_1: Eco RI;  
 Site\_2: Xho I; Lower temperature 50 C / hour from 22 to  
 120C; bring to 50 in 1 hour from 120C. Leave at 50C 2 days  
 , photoperiod 16 hours. Light intensity was 125 uE-1.  
 library prepared by in vivo mass excision from amplified  
 library."  
 BASE COUNT 118 a 263 c 213 g 80 t 36 others  
 ORIGIN

Query Match 20.3%; Score 43.4; DB 12; Length 710;  
 Best Local Similarity 55.6%; Pred. No. 8.5;  
 Matches 80: Conservative 1; Mismatches 63; Indels 0; Gaps 0;

OY 18 GATCTGGCCCGGCTGGTCTGCTGACAGTGAACAAGAGCTGGCCCTGGCGGCAACCAAT 77  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 DB 355 GAGAGGAGCCNCATGCGCTGTTCTATGAGAGGCGGCGCATCCAGAMCTGGCCAGGCTCCC 414  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 OY 78 CCGCTCAGAGTGAACCAACCAACCTCTACGAGAGAGAGCGTGGCCACCATCTGAGTAC 137  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 DB 415 ATGAGAGAGAGTGAACCGCCGCAAGCTTCACCAAGAGCAGCAGCTGACGTGCTCAACTTG 474  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 OY 138 CTGTGCGCTCTGACGAGGCGGCAAG 161  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 DB 475 CTGCGCTCTTCTCCACCGCGGCAAG 498  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 14  
 AY109418 1856 bp mRNA linear HTC 25-MAY-2002  
 LOCUS  
 DEFINITION Zea mays CL2386\_1 mRNA sequence.  
 AY109418  
 ACCESSION  
 VERSION AY109418.1 GI:21213135  
 KEYWORDS HTC.  
 SOURCE Zea mays.  
 ORGANISM Zea mays.

REFERENCE  
 AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC  
 clade; Panicoidae; Andropogoneae; Zea.  
 HAINES, C.F., DOLAN, M., MIAO, G.H., VOGEL, J.M., WHITSITT, M.S.,  
 ARTHUR, L.W., HANLEY, M., MORGANTE, M., and TINKER, S.V.  
 TITILE Maize Mapping Project/Dupont Consensus Sequences for Design of  
 JOURNAL Overgo Probes  
 Unpublished (2002)  
 2 (bases 1 to 1856)  
 COE, E.C.  
 REFERENCE Direct Submission  
 Submitted (25-APR-2002) Maize Mapping Project, University of  
 Missouri, Columbia, MO 65211, USA  
 JOURNAL Location/Qualifiers  
 1.1856

/organism="Zea mays"  
 /db\_xref="MaizeDB:630950"  
 /db\_xref="taxon:4577"  
 /clone="CL2386\_1"  
 /clone\_1lb="Maize Mapping Project/Dupont Consensus  
 Library"  
 /note="this sequence is part of a project of EST  
 assemblies resulting from the application of public  
 contigs to seed Dupont contigs; this resource was  
 assembled by Dupont as part of a collaboration for the

overgo addressing of BACS in conjunction with the Maize Mapping Project  
BASE COUNT 320 a 548 c 585 g 281 t 122 others  
ORIGIN

Query Match 20.2% Score 43.2; DB 11: Length 1856;  
Best Local Similarity 56.2% Pred. No. 11;  
Matches 81: Conservative 0; Mismatches 63; Indels 0; Gaps 0;

OY 18 GATCGGCCCCGCTGGGTGCTACAGAGTGAACAGAGCTGGGCGGACCAATCGGCTC 77  
DB 107 GAGAGGAGGCCCATGCGCTGTTTCATGAGGCGGCGCATCCAGACCTGGGCGAGGTCGCC 166  
OY 78 CCGGCTGAGTGACGACACCACTCAGGAGAGAGCTGCTGCCACATGAGTAC 137  
DB 167 ATGGACGAGATCGAGCGGCTCAAGCTCAGCAGAGAGAGCTGCGCTCCGCTCAAGTGG 226  
OY 138 CTGCTGGCGCTGCACGAGGCGCAG 161  
DB 227 CTCGGCCTCTCCACGCCGCGCAG 250

RESULT 15  
BE754521 558 bp mRNA linear EST 25-APR-2001  
LOCUS 208060 MARC 2BOV Bos taurus cDNA 5', mRNA sequence.  
DEFINITION BE754521  
ACCESSION BE754521  
VERSION BE754521.1 GI:10168513  
KEYWORDS EST.  
SOURCE cow.  
ORGANISM Bos taurus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
Bovidae; Bovinae; Bos.

REFERENCE 1 (bases 1 to 558)  
AUTHORS Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,  
Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett,  
G.L., Heaton,M.P., Laegreid,W.M., Rohrer,G.A., Chitko-Mckown,C.G.,  
Pettea,G., Holt,I., Karameycheva,S., Liang,F., Quackenbush,J. and  
Keefe,J.W.  
TITLE Sequence evaluation of four pooled-tissue normalized bovine cDNA  
libraries and construction of a gene index for cattle  
JOURNAL Genome Res. 11 (4), 626-630 (2001)  
MEDLINE 21180013  
COMMENT Contact: Smith TPL  
USDA, ARS, US Meat Animal Research Center  
PO Box 166, Clay Center, NE 68933-0166, USA  
Tel: 402 762 4366  
Fax: 402 762 4390  
Email: smitht@mail.marc.usda.gov

Single pass sequencing. Bases called and alt-trimmed with phred  
v0.980904.e. Vector identified by cross\_match with the -minscore 18  
and -mismatch 12 options.  
PCR Primers  
FORWARD: AGGAACAGCTATGACCAT  
BACKWARD: GTTTCCTCAGTCAGCAG  
Plate: 54 row: M column: 18  
Seq primer: ATTTAGTGACACTATAG.  
FEATURES  
source  
1..558  
/organism="Bos taurus"  
/db\_xref="taxon:9913"  
/clone\_lib="MARC 2BOV"  
/tissue\_type="pooled"  
/lab\_host="DH10B"  
/note="Vector: PCMV SPORT6; Site\_1: NotI; Site\_2: SalI;  
Library made from pooled tissue from testis, thymus,  
semitendinosus muscle, longissimus muscle, pancreas,  
adrenal, and endometrium."

BASE COUNT 113 a 166 c 172 g 107 t  
ORIGIN

Query Match 20.1% Score 43; DB 12: Length 558;

Best Local Similarity 54.0% Pred. No. 10;  
Matches 88: Conservative 0; Mismatches 75; Indels 0; Gaps 0;

OY 25 CCCCCTGGGCTGCTACAGGTGAACAAGAGCTGGGCGGACCAATCGGCTC 84  
DB 198 CCATCATGGGGCAGTTCGACACCCCAATATCATCCGCGGAGGGCGTGCACCAACA 257  
OY 85 AGTGACCAACCAACCTCAGCGAGAGAGCTGCTCGCACATGAGTACTGCTGC 144  
DB 258 GCGTCCCGTATGATCTCTCAGCAGTTCATGAGAGAGAGTGGCGCTGCACTCTCTGC 317  
OY 145 GCTGCAGAGAGGGCAGACCAAGATGACGCGCGCGGCGCT 187  
DB 318 GCGTGAAGAGAGGGCAGATTACCGTATCCAGCTCTGTTGGCAT 360

Search completed: November 13, 2002, 04:00:40  
Job time : 1154.63 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 12, 2002, 16:05:40 ; Search time 21.7711 Seconds

(without alignments)  
3487.380 Million cell updates/sec

Title: US-09-697-123b-21

Perfect score: 214  
Sequence: 1 tcaagagagacgcacgacat.....ccggtgatgtgacacacat 214

Scoring table: IDENTITY-NUC  
Gapop 10.0 , Gapext 1.0

Searched: 320260 seqs, 177392727 residues

Total number of hits satisfying chosen parameters: 640520

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

Published Applications -NA:\*

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/2/pubpna/PCRT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/2/pubpna/PCRTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*
- 12: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*
- 13: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*
- 14: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	80.4	37.6	5096	10 US-09-984-711-5	Sequence 5, Appl1
2	80.4	37.6	5099	9 US-10-075-460-5	Sequence 5, Appl1
3	80.4	37.6	5099	10 US-09-887-052-1	Sequence 1, Appl1
4	80.4	37.6	5099	10 US-09-887-052-3	Sequence 3, Appl1
5	80.4	37.6	5099	10 US-09-887-052-5	Sequence 5, Appl1
6	41	19.2	399	10 US-09-878-574-3598	Sequence 3598, Ap
7	38.6	18.0	392	10 US-09-861-893-29	Sequence 29, Appl
8	38.6	18.0	431	10 US-09-954-456-1034	Sequence 1034, Ap
9	38.2	17.9	1818	10 US-09-815-242-7989	Sequence 7989, Ap
10	37.2	17.4	250	10 US-09-815-242-7989	Sequence 229, App
11	37.2	17.4	2383	10 US-09-764-853-229	Sequence 11, Appl
12	37	17.3	879	10 US-09-825-302-317	Sequence 3, Appl1
13	37	17.3	879	10 US-09-861-289-11	Sequence 11, Appl
14	36.8	17.2	13613	10 US-09-861-289-3	Sequence 7681, Ap
15	36.6	17.1	855	10 US-09-815-242-7681	Sequence 33210, A
16	36	16.8	350	10 US-09-864-761-33210	Sequence 313, App
17	35.6	16.6	4252	10 US-09-815-242-4133	Sequence 4133, Ap
18	35.6	16.6	1680	10 US-09-820-721A-2	Sequence 2, Appl1
19	35.4	16.5	409	10 US-09-878-574-3856	Sequence 3856, Ap

20	35.4	16.5	3057	10 US-09-815-242-4131	Sequence 4131, Ap
21	35.2	16.4	2712	10 US-09-748-033-4	Sequence 4, Appl1
22	35.2	16.4	3435	10 US-09-803-670-1	Sequence 1, Appl1
23	35.2	16.4	8868	10 US-09-803-670-3	Sequence 3, Appl1
24	35	16.4	3945	10 US-09-921-771-4	Sequence 4, Appl1
25	35	16.4	3969	10 US-09-982-610-23	Sequence 23, Appl
26	35	16.4	7419	10 US-09-815-242-4009	Sequence 4009, Ap
27	34.8	16.3	403	10 US-09-878-574-105	Sequence 2105, Ap
28	34.8	16.3	1101	10 US-09-815-242-7943	Sequence 7943, Ap
29	34.8	16.3	1416	10 US-09-815-242-7943	Sequence 4160, Ap
30	34.8	16.3	1476	10 US-09-997-664-99	Sequence 99, Appl
31	34.8	16.3	6491	10 US-09-732-680A-1	Sequence 1, Appl1
32	34.6	16.2	1539	10 US-09-772-304A-1	Sequence 1, Appl1
33	34.6	16.2	1914	10 US-09-815-242-7960	Sequence 7960, Ap
34	34.6	16.2	2090	10 US-09-822-887-3	Sequence 105, App
35	34.6	16.2	2142	10 US-09-822-887-3	Sequence 367, App
36	34.6	16.2	2493	10 US-09-822-887-1	Sequence 1, Appl1
37	34.6	16.2	4826	10 US-09-772-304A-1	Sequence 1, Appl1
38	34.4	16.1	2329	10 US-09-816-828-9	Sequence 9, Appl1
39	34.2	16.0	1467	10 US-09-887-576-777	Sequence 777, App
40	34	15.9	480	12 US-10-052-586-367	Sequence 367, App
41	33.8	15.8	11575	10 US-09-938-956-5	Sequence 5, Appl1
42	33.8	15.8	12241	12 US-10-074-279-10	Sequence 10, Appl1
43	33.8	15.8	13737	9 US-10-074-279-10	Sequence 4, Appl1
44	33.8	15.8	14446	9 US-09-815-242-7949	Sequence 7949, Ap
45	33.6	15.7	801	10 US-09-815-242-7949	

#### ALIGNMENTS

RESULT 1  
US-09-984-711-5  
Sequence 5, Application US/09984711  
Patent No. US20020119549A1  
GENERAL INFORMATION:  
APPLICANT: MOECKEL, Bettina  
APPLICANT: BATHÉ, Brigitte  
APPLICANT: STEPHAN, Hans  
APPLICANT: KREUTZER, Caroline  
APPLICANT: HERMANN, Thomas  
APPLICANT: PFEFFERLE, Walter  
APPLICANT: BINDER, Michael  
TITLE OF INVENTION: NUCLEOTIDE SEQUENCES WHICH CODE FOR THE TP51 GENE  
FILE REFERENCE: 204209US0  
CURRENT APPLICATION NUMBER: US/09/984, 711  
CURRENT FILING DATE: 2001-10-31  
PRIOR APPLICATION NUMBER: DE10108230.9  
NUMBER OF SEQ ID NOS: 6  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 5  
LENGTH: 5096  
TYPE: DNA  
ORGANISM: Corynebacterium glutamicum  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (702)..(4196)  
OTHER INFORMATION:  
US-09-984-711-5  
Query Match  
Best Local Similarity 66.08; Pred. No. 1.3e-11;  
Matches 136; Conservative 0; Mismatches 61; Indels 9; Gaps 1;  
DB 1578 AAGCGCTACGACCTGCGGTGCTGTACAGATCAACCGCAAGCTGGCTTGGT 1637  
DB 69 GGCACCAATCGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGT 128  
DB 1638 GGCACCAATCGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGT 1688

OY 129 ATCGAGTACCTGGTGGCCCTGCACGAGGCCAGACCAGCATGACCGCCCGGCGGCTC 188  
|||||  
Db 1689 ATCGAGTACCTGGTGGCTGTGCACGAGGTGAGCGGTCATGACTTCTCCAAATGGTGAA 1748  
OY 189 GAGTGCCGGTGGATGTGACGACAT 214  
|||  
Db 1749 GAGATCCCACTGACGACCATGACAT 1774

## RESULT 2

US-10-075-460-5  
; Sequence 5, Application US/10075460  
; Patent No. US2002015557A1  
; GENERAL INFORMATION:  
; APPLICANT: MOCKEL, BETTINA  
; APPLICANT: BATHE, BRIGITTE  
; APPLICANT: HANS, STEFAN  
; APPLICANT: KREUTZER, CAROLINE  
; APPLICANT: HERMANN, THOMAS  
; APPLICANT: PFEFFERLE, WALTER  
; APPLICANT: BINDER, MICHAEL  
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES WHICH CODE FOR THE rpsL GENE  
; FILE REFERENCE: 218472050X  
; CURRENT APPLICATION NUMBER: US/10/075.460  
; PRIOR FILING DATE: 2002-02-15  
; PRIOR APPLICATION NUMBER: DE 10107230.9  
; PRIOR FILING DATE: 2001-02-16  
; PRIOR APPLICATION NUMBER: DE 10162386.0  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 5  
; LENGTH: 5099  
; TYPE: DNA  
; ORGANISM: Corynebacterium glutamicum  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (702)..(4196)  
; OTHER INFORMATION:  
US-10-075-460-5

Query Match 37.6%; Score 80.4; DB 9; Length 5099;  
Best Local Similarity 66.0%; Pred. No. 1.3e-11;

Matches 136; Conservative 0; Mismatches 61; Indels 9; Gaps 1;

OY 9 AAGCGTACGATCTGGCCCGGCGGTGCGGTACAGGTGACGAAGAAGCTGGGCTGGGC 68  
|||||  
Db 1578 AAGCGTACGATCTGGCTGCTGCTGCTGCTTACAGATCAACCGCAAGCTGCGCTTGGT 1637  
OY 69 GGCACCAATCCGGCTCAGGTGACCCACCAACCCTCAGCGAGGAAGCTGCGCCACC 128  
|||||  
Db 1638 GGCACCAATCCGGCTGTTGATGACT-----CTTACTGAAGAGAGCATCCCAACACC 1688  
OY 129 ATCGAGTACCTGGTGGCCCTGCACGAGGCCAGACCAGCATGACCGCCCGGCGGCTC 188  
|||||  
Db 1689 ATCGAGTACCTGGTGGCTGTGCACGAGGTGAGCGGCTGATCTTCCAAATGGTGAA 1748  
OY 189 GAGTGCCGGTGGATGTGACGACAT 214  
|||  
Db 1749 GAGATCCCACTGACGACCATGACAT 1774

## RESULT 3

US-09-887-052-1  
; Sequence 1, Application US/09887052  
; Patent No. US20020119537A1  
; GENERAL INFORMATION:  
; APPLICANT: MOCKEL, Bettina  
; APPLICANT: BATHE, Brigitte  
; APPLICANT: HERMANN, Thomas  
; APPLICANT: PFEFFERLE, Walter  
; APPLICANT: BINDER, Michael  
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES WHICH CODE FOR THE rpoB GENE

; FILE REFERENCE: 204212050X  
; CURRENT APPLICATION NUMBER: US/09/887.052  
; CURRENT FILING DATE: 2001-06-25  
; PRIOR APPLICATION NUMBER: DE10107229.5  
; PRIOR FILING DATE: 2001-02-16  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 1  
; LENGTH: 5099  
; TYPE: DNA  
; ORGANISM: Corynebacterium glutamicum  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (702)..(4196)  
US-09-887-052-1

Query Match 37.6%; Score 80.4; DB 10; Length 5099;  
Best Local Similarity 66.0%; Pred. No. 1.3e-11;  
Matches 136; Conservative 0; Mismatches 61; Indels 9; Gaps 1;

OY 9 AAGCGTACGATCTGGCCCGGCGGTGCGGTACAGGTGACGAAGAAGCTGGGCTGGGC 68  
|||||  
Db 1578 AAGCGTACGATCTGGCTGCTGCTGCTGCTTACAGATCAACCGCAAGCTGCGCTTGGT 1637  
OY 69 GGCACCAATCCGGCTCAGGTGACCCACCAACCCTCAGCGAGGAAGCTGCGCCACC 128  
|||||  
Db 1638 GGCACCAATCCGGCTGTTGATGACT-----CTTACTGAAGAGAGCATCCCAACACC 1688  
OY 129 ATCGAGTACCTGGTGGCCCTGCACGAGGCCAGACCAGATGACCGCCCGGCGGCTC 188  
|||||  
Db 1689 ATCGAGTACCTGGTGGCTGTGCACGAGGTGAGCGGCTGATCTTCCAAATGGTGAA 1748  
OY 189 GAGTGCCGGTGGATGTGACGACAT 214  
|||  
Db 1749 GAGATCCCACTGACGACCATGACAT 1774

## RESULT 4

US-09-887-052-3  
; Sequence 3, Application US/09887052  
; Patent No. US20020119537A1  
; GENERAL INFORMATION:  
; APPLICANT: MOCKEL, Bettina  
; APPLICANT: BATHE, Brigitte  
; APPLICANT: HERMANN, Thomas  
; APPLICANT: PFEFFERLE, Walter  
; APPLICANT: BINDER, Michael  
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES WHICH CODE FOR THE rpoB GENE  
; FILE REFERENCE: 204212050X  
; CURRENT APPLICATION NUMBER: US/09/887.052  
; CURRENT FILING DATE: 2001-06-25  
; PRIOR APPLICATION NUMBER: DE10107229.5  
; PRIOR FILING DATE: 2001-02-16  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 3  
; LENGTH: 5099  
; TYPE: DNA  
; ORGANISM: Corynebacterium glutamicum  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (702)..(4196)  
US-09-887-052-3

Query Match 37.6%; Score 80.4; DB 10; Length 5099;  
Best Local Similarity 66.0%; Pred. No. 1.3e-11;  
Matches 136; Conservative 0; Mismatches 61; Indels 9; Gaps 1;

OY 9 AAGCGTACGATCTGGCCCGGCGGTGCGGTACAGGTGACGAAGAAGCTGGGCTGGGC 68  
|||||  
Db 1578 AAGCGTACGATCTGGCTGCTGCTGCTGCTTACAGATCAACCGCAAGCTGCGCTTGGT 1637  
OY 69 GGCACCAATCCGGCTCAGGTGACCCACCAACCCTCAGCGAGGAAGCTGCGCCACC 128



```
Db 1638 GCGGACACGATGTTGATGACT-----CTTACGAAAGAGACGATCGCAACACC 1688
QY 129 ATGAGTACTGTTGGCTGACGAGGCGCCAGCAGATGACCGCCCGCGCGCTC 188
Db 1689 ATGAGTACTGTTGGCTGACGAGGCGCGCTGATGACTTCTCAAAATGGTGA 1748
QY 189 GAGTCCCGGTGATGAGCAGCAT 214
Db 1749 GAGATCCCAAGTCGAGACGATGACAT 1774
```

## RESULT 5

```
US-09-887-052-5
; Sequence 5, Application US/09887052
; Patent No. US20020119537A1
; GENERAL INFORMATION:
; APPLICANT: MOECKEL, Bettina
; APPLICANT: BATHÉ, Brigitte
; APPLICANT: HERMANN, Thomas
; APPLICANT: PEPPERLE, Walter
; APPLICANT: BINDER, Michael
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES WHICH CODE FOR THE TP0B GENE
; FILE REFERENCE: 204212US0X
; CURRENT APPLICATION NUMBER: US/09/887,052
; PRIOR FILING DATE: 2001-08-25
; PRIOR APPLICATION NUMBER: DE10107229.5
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
; LENGTH: 5099
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (702)..(4196)
US-09-887-052-5
```

```
Query Match 37.6%; Score 80.4; DB 10; Length 5099;
Best Local Similarity 66.0%; Pred. No. 1.3e-11;
Matches 136; Conservative 0; Mismatches 61; Indels 9; Gaps 1;
```

```
QY 9 AAGCCCTACGATCTGGCCGCGTGGTGGTACAGAGTGAACAAGACCTGGCGGC 68
Db 1578 AAGCCCTACGATCTGGCCGCGTGGTGGTACAGATCAAGACGAGCTGGCGGC 1637
QY 69 GCGACCAATCCGCTCAGGTGACCCACCACCTCAGGAGAGAGCGTGGCGGAC 128
Db 1638 GCGACCAATCCGCTCAGGTGACCCACCACCTCAGGAGAGAGCGTGGCGGAC 1688
QY 129 ATCGAGTACCTGGTGGCCCTGACAGAGGCGCCAGACGATGACCGCCCGCGCGCTC 188
Db 1689 ATCGAGTACCTGGTGGCCCTGACAGAGGCGCGTCAATGCTGCTCAAAATGGTGA 1748
QY 189 GAGTCCCGGTGATGAGCAGCAT 214
Db 1749 GAGATCCCAAGTCGAGACGATGACAT 1774
```

## RESULT 6

```
US-09-878-574-3598
; Sequence 3598, Application US/09878574
; Patent No. US20020110548A1
; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Thompson, Michael D.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(15401)B
; CURRENT APPLICATION NUMBER: US/09/878,574
; CURRENT FILING DATE: 2001-12-21
```

```
; PRIOR APPLICATION NUMBER: 09/333,535
; PRIOR FILING DATE: 1999-06-14
; NUMBER OF SEQ ID NOS: 15775
; SEQ ID NO 3598
; LENGTH: 392
; TYPE: DNA
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: LIB3028-008-Q1-B1-C6
US-09-878-574-3598
```

```
Query Match 19.2%; Score 41; DB 10; Length 392;
Best Local Similarity 56.2%; Pred. No. 0.051;
Matches 77; Conservative 0; Mismatches 60; Indels 0; Gaps 0;
```

```
QY 76 ATCCGCTCAGGTGACCAACACACCTGACGAGGAGGAAGAGTGTGCGCCACCATGAGT 135
Db 2 ACCTCGCCAGGCTTCCGCTCATCGACCGCCGACACGAGGCTGTGACGCGCGCAGG 61
QY 136 ACCTGCTGCGCTGACGAGGCGCCAGACCATGATGACCGCCCGCGCGCTGAGGTGC 195
Db 62 ACCTCGAGGCTCTCTCAGGCTGCGCGCTGCGCGCTGCGCGCGCGCGCGCGCTGAC 121
QY 196 CGGTGATGTGAGCAGC 212
Db 122 TGCTCGGAGGCTGAC 138
```

## RESULT 7

```
US-09-861-893-29
; Sequence 29, Application US/09861893
; Patent No. US20020045257A1
; GENERAL INFORMATION:
; APPLICANT: Feinberg, Andrew
; APPLICANT: Strichman-Almashanu, Liora
; APPLICANT: Jiang, Shan
; TITLE OF INVENTION: METHODS FOR ASSAYING GENE IMPRINTING AND
; FILE REFERENCE: 01107,00128
; CURRENT APPLICATION NUMBER: US/09/861,893
; PRIOR FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: 60/206,158
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/206,161
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 29
; LENGTH: 399
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-861-893-29
```

```
Query Match 18.0%; Score 38.6; DB 10; Length 399;
Best Local Similarity 52.1%; Pred. No. 0.2;
Matches 86; Conservative 0; Mismatches 79; Indels 0; Gaps 0;
```

```
QY 36 CGGTCAAGGTGAACAAGAGTGGCGTGGGCGGACCAATCCGCTAGTGAACACC 95
Db 224 CAGACCAAGTGTCTCAACATCGCAACGAGACTGATCAAGGTGGCGCTGTCTCAAC 283
QY 96 ACCACCTTACCGAGAGAGAGTGTGCGCACCATGAGTACTGTTGCGCTGACGAG 155
Db 284 AACGCTTCTACTGTGAGAACTGTGACTTACACATGAGAGGAGAGACGACTACTTC 343
QY 156 GCGCAGACGAGTGAACCGCCCGCGCGCGCTGAGAGTCCCGTG 200
Db 344 ATCAAGACACACGCGCCGAGAGCGACTGGCAGCGCTGCGGTG 388
```

## RESULT 8

```
US-09-954-456-1034/C
; Sequence 1034, Application US/09954456
; Patent No. US20020115057A1
```

```

: GENERAL INFORMATION:
: APPLICANT: Young, Paul
: TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cand
: TITLE OF INVENTION: Sets
: FILE REFERENCE: 689290-76
: CURRENT APPLICATION NUMBER: US/09/954,456
: CURRENT FILING DATE: 2001-09-18
: PRIOR APPLICATION NUMBER: US/60/233,617
: PRIOR FILING DATE: 2000-09-18
: PRIOR APPLICATION NUMBER: US/60/234,052
: PRIOR FILING DATE: 2000-09-20
: PRIOR APPLICATION NUMBER: US/60/234,923
: PRIOR FILING DATE: 2000-09-25
: PRIOR APPLICATION NUMBER: US/60/235,134
: PRIOR FILING DATE: 2000-09-25
: PRIOR APPLICATION NUMBER: US/60/235,637
: PRIOR FILING DATE: 2000-09-26
: PRIOR APPLICATION NUMBER: US/60/235,638
: PRIOR FILING DATE: 2000-09-26
: PRIOR APPLICATION NUMBER: US/60/235,711
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: US/60/235,720
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: US/60/235,840
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: US/60/235,863
: NUMBER OF SEQ ID NOS: 2276
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 1034
: LENGTH: 431
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-954-456-1034

Query Match          18.0%: Score 38.6; DB 10; Length 431;
Best Local Similarity 52.1%: Pred. No. 0.2;
Matches 86; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

QY 36 CGGTACAGGTGAACAAGAGCTGGGCGTGGCGGCGACCAATCCGCTCAGGTGACGACC 95
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 177 CAGACCAACGTGCTCAACATCCCAAGAGAGACTGCATCAAGTGGCGCGCTGCTCAC 118
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 96 ACCACCTCAGCAGAGAAAGCGTCGTCCGCCATCGATCGATCTGCTCGCTCAGCAG 155
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 117 AACGCTTTTACTCTGAGAACCTGCATCTTCAACCATCGAGGCGCAAGACACGCACTACTTC 58
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 156 GGGCAGACGACGATGACCGCCCGCGGCGCTGCGAGTGGCCGTTG 200
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 57 ATCAAGACCACGACCGCCGAGAGCGACCTGGGCGACGCTGCGGTTG 13

RESULT 9
US-09-815-242-7989
: Sequence 7989, Application US/09815242
: Patent No. US20020061569A1
: GENERAL INFORMATION:
: APPLICANT: Haselbeck, Robert
: APPLICANT: Ohlsen, Karl L.
: APPLICANT: Zyskind, Judith W.
: APPLICANT: Wall, Daniel
: APPLICANT: Trawick, John D.
: APPLICANT: Carr, Grant J.
: APPLICANT: Yamamoto, Robert T.
: APPLICANT: Xu, H. Howard
: TITLE OF INVENTION: Identification of Essential Genes in
: TITLE OF INVENTION: Prokaryotes
: FILE REFERENCE: ELITRA.011A
: CURRENT APPLICATION NUMBER: US/09/815,242
: CURRENT FILING DATE: 2001-03-21
: PRIOR APPLICATION NUMBER: 60/191,078
: PRIOR FILING DATE: 2000-03-21
: PRIOR APPLICATION NUMBER: 60/206,848
```

```

: PRIOR FILING DATE: 2000-05-23
: PRIOR APPLICATION NUMBER: 60/207,727
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: 60/242,578
: PRIOR FILING DATE: 2000-10-23
: PRIOR APPLICATION NUMBER: 60/253,625
: PRIOR FILING DATE: 2000-11-27
: PRIOR APPLICATION NUMBER: 60/257,931
: PRIOR FILING DATE: 2000-12-22
: PRIOR APPLICATION NUMBER: 60/269,308
: PRIOR FILING DATE: 2001-02-16
: NUMBER OF SEQ ID NOS: 14110
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 7989
: LENGTH: 1818
: TYPE: DNA
: ORGANISM: Pseudomonas aeruginosa
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (1)...(1818)
US-09-815-242-7989

Query Match          17.9%: Score 38.2; DB 10; Length 1818;
Best Local Similarity 53.3%: Pred. No. 0.29;
Matches 104; Conservative 0; Mismatches 88; Indels 3; Gaps 1;
```

```

QY 23 GGGCCGCGTGGGTGCGGTACAGGTGAACAAGAGCTGGGCGTGGCGCACCAATCCGCG 82
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 354 GGGCGGCGTGGCGTCCGATGCTGTGTAAGATTCAGACCGCCGGCGCGCTCGGA 413
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 83 TCAGGTGACACACCAACCTCAGCAG--GAGAGCTGCTGGCCACCATCGATCT 139
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 414 CTGGGTATCGACAGCATCTTGCACCTTTCGACAACTCGGCGGACCGACGAGCACT 473
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 140 GGTGCCCTGTCAGAGGCGACAGCAGATGACCGCCCGCGCGCTGAGGTGCCGCT 199
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 474 CGACTTCGCGATGCTGTACGCGCGCCCTGACCGCATCCGCGCTGACACGAGAA 533
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 200 GATGTGAGCACT 214
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 534 GATGAGCAACAT 548
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 10
US-09-764-853-229/c
: Sequence 229, Application US/09764853
: Patent No. US20020090672A1
: GENERAL INFORMATION:
: APPLICANT: Rosen et al.
: TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
: FILE REFERENCE: P1206
: CURRENT APPLICATION NUMBER: US/09/764,853
: CURRENT FILING DATE: 2001-01-17
: PRIOR APPLICATION data removed - consult PALM or file wrapper
: NUMBER OF SEQ ID NOS: 939
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 229
: LENGTH: 250
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-764-853-229

Query Match          17.4%: Score 37.2; DB 10; Length 250;
Best Local Similarity 56.9%: Pred. No. 0.42;
Matches 66; Conservative 1; Mismatches 49; Indels 0; Gaps 0;
```

```

QY 13 GCTACGATCTGCGCCGCGTGGCTGCTACAGGTGAACAAGAGCTGGCGCGCGCA 72
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 125 GCTCACACTGAGAGCTTGCGGTGGAGTGAAGTGAACATTCGAGGCGGAAGCGC 66
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 73 CCAATCCGGCTCAGGTGACCAACCAACCCCTCACCGAGGAAGAGCTGTGCGCAC 128
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 65 CCACTTCCGAGAGCTCCACTCGACACCCGCTGGAAGCGCTCTCCGCGCC 10
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
RESULT 11
US-09-925-302-317
: Sequence 317, Application US/09925302
: Patent No. US20020044941A1
: GENERAL INFORMATION:
: APPLICANT: Rosen et al.
: TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
: FILE REFERENCE: PA104
: CURRENT APPLICATION NUMBER: US/09/925.302
: CURRENT FILING DATE: 2001-08-10
: PRIOR APPLICATION NUMBER: PCT/US00/05918
: PRIOR FILING DATE: 2000-03-08
: PRIOR APPLICATION NUMBER: 60/124,270
: PRIOR FILING DATE: 1999-03-12
: NUMBER OF SEQ ID NOS: 896
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 317
: LENGTH: 2383
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-925-302-317

Query Match
Best Local Similarity 17.4%; Score 37.2; DB 10; Length 2383;
Matches 96; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

OY 8 GAAGCGCTACGATCTGGCGCCGCTGGGTGCAAGTGACAGAGAGCTGGGCTGGG 67
Db 838 GCACCCAGAGAGCGGGTCTATCTGAACTCAGCAAGTCTCTGAGAGAGTG 897
OY 68 CGGCACCAATCGGCTCAGGTGACCAACCACTCAGGAGAGAGCTGTCGCCAC 127
Db 898 CGAGCGCGGAGTGCAGCGCCGACTGCAGCAACCCCTCACCCTGACAGAGTCCCGCAT 957
OY 128 CATGAGTACCTGTGGCGCCCTGCAGAGGCGCAGACATACCGCCCGCGCGCT 187
Db 958 CTTGGAACCTCGAGAGCAAGTACTATGAGAGTACAGATGTCCGACCTGTGACCT 1017
OY 188 CGAGTCCCGGTG 201
Db 1018 TGTGTGTGGCATCG 1031

RESULT 12
US-09-861-289-11
: Sequence 11, Application US/09861289
: Patent No. US20020110897A1
: GENERAL INFORMATION:
: APPLICANT: Sherman, D.H.
: APPLICANT: Liu, H.
: APPLICANT: Xue, Y.
: APPLICANT: Zhao, L.
: TITLE OF INVENTION: DNA encoding methymycin and pikromycin
: FILE REFERENCE: 600.438US1
: CURRENT APPLICATION NUMBER: US/09/861,289
: CURRENT FILING DATE: 2001-05-18
: PRIOR APPLICATION NUMBER: 09/105,537
: PRIOR FILING DATE: 1998-06-26
: NUMBER OF SEQ ID NOS: 43
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 11
: LENGTH: 879
: TYPE: DNA
: ORGANISM: Streptomyces venezuelae
US-09-861-289-11

Query Match
Best Local Similarity 17.3%; Score 37; DB 10; Length 879;
Matches 70; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

OY 87 GTGACCAACCACTCAGGAGAGAGTGTGCGCACCATCGATGCTGTGCGC 146
```

```
Db 508 GTACCGGCTCTACTCTACAGACAGAGCTGTCGACATCGCAAGAACATCGGCCC 567
OY 147 CTGCAGAGGCGACACAGATGACCGCCCGCGCGCTGAGGTGCGCGTGTGATG 206
Db 568 TCGCCCGCGCGGAGCTGAGATCAACGAGCTCAACCGCTTACTGTGAGCGGCGCG 627
OY 207 GACGA 211
Db 628 GCCGA 632

RESULT 13
US-09-861-289-3/c
: Sequence 3, Application US/09861289
: Patent No. US20020110897A1
: GENERAL INFORMATION:
: APPLICANT: Sherman, D.H.
: APPLICANT: Liu, H.
: APPLICANT: Xue, Y.
: APPLICANT: Zhao, L.
: TITLE OF INVENTION: DNA encoding methymycin and pikromycin
: FILE REFERENCE: 600.438US1
: CURRENT APPLICATION NUMBER: US/09/861,289
: CURRENT FILING DATE: 2001-05-18
: PRIOR APPLICATION NUMBER: 09/105,537
: PRIOR FILING DATE: 1998-06-26
: NUMBER OF SEQ ID NOS: 43
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 3
: LENGTH: 13613
: TYPE: DNA
: ORGANISM: Streptomyces venezuelae
US-09-861-289-3

Query Match
Best Local Similarity 17.3%; Score 37; DB 10; Length 13613;
Matches 70; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

OY 87 GTGACCAACCACTCAGGAGAGAGTGTGCGCACCATCGATGCTGTGCGC 146
Db 9482 GTCAACCGGCTCTACTCTACAGACAGAGCTGTCGACATGCGCAAGAACATCGGCCC 9423
OY 147 CTGCAGAGGCGACACAGATGACCGCCCGCGCGCTGAGGTGCGCGTGTGATG 206
Db 9422 TCGCCCGCGCGGAGCTGAGATCACCAGATCAACCGCTTACTGTGAGCGGCGCG 9363
OY 207 GACGA 211
Db 9362 GCCGA 9358

RESULT 14
US-09-815-242-7681/c
: Sequence 7681, Application US/09815242
: Patent No. US20020061569A1
: GENERAL INFORMATION:
: APPLICANT: Haselbeck, Robert
: APPLICANT: Ohlson, Karl L.
: APPLICANT: Zyskind, Judith W.
: APPLICANT: Wall, Daniel
: APPLICANT: Trawick, John D.
: APPLICANT: Carr, Grant J.
: APPLICANT: Yamamoto, Robert T.
: APPLICANT: Xu, H. Howard
: TITLE OF INVENTION: Identification of Essential Genes in
: FILE REFERENCE: ELITRA.011A
: CURRENT APPLICATION NUMBER: US/09/815,242
: CURRENT FILING DATE: 2001-03-21
: PRIOR APPLICATION NUMBER: 60/191,078
: PRIOR FILING DATE: 2000-03-21
: PRIOR APPLICATION NUMBER: 60/206,848
```

```

; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: fastseq for windows Version 4.0
; SEQ ID NO 7681
; LENGTH: 855
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1) ... (855)
US-09-815-242-7681
```

```
Query Match          17.2% Score 36.8; DB 10; Length 855;
Best Local Similarity 50.0%; Pred. No. 0.59;
Matches 92; Conservative 0; Mismatches 92; Indels 0; Gaps 0;
```

```

OY 28 GCGTGGCTCGGTACAGGTACACAGAGCTGGGCGGCGGCAACCAATCCGCTCAGG 87
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 500 GCGTGGCGCGCGCGCGCGCGGTGTCGACGACAGAGCAGCGGCTCGAGCGGCTCGGCGAGC 441
OY 88 TGACACACACACACCTCAGCAGAGAGAGAGAGTGTGCGGCGGCAACATCGAGTCTGGTGGCC 147
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 440 AGGCGATGGGCGACCTGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 381
OY 148 TGCACGAGGCGCGACACAGATGACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 207
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 380 GCGCGCTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 321
OY 208 ACGA 211
    |||
DB 320 CCGA 317
```

```

RESULT 15
US-09-864-761-33210/c
; Sequence 33210, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
```

```

; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 33210
; LENGTH: 350
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC005973.2
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL - 0.88
; OTHER INFORMATION: EST HUMAN HIT: AA454168.1, EVALUE 2.20e-01
; OTHER INFORMATION: NT HIT: AF10599.1, EVALUE 1.00e-127
; OTHER INFORMATION: SWISSPROT HIT: Q04844, EVALUE 3.00e-22
US-09-864-761-33210
```

```
Query Match          17.1% Score 36.6; DB 10; Length 350;
Best Local Similarity 50.9%; Pred. No. 0.61;
Matches 87; Conservative 0; Mismatches 84; Indels 0; Gaps 0;
```

```

OY 24 GCGCGGTGGTGTGCTACAGGTGACAGAGAGCTGGGCGTGGCGGCGGCGGCGGCGGCGGCT 83
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 231 GCCCAGAGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 172
OY 84 CAGGTGACACACACACCTCAGCAGAGAGAGAGAGCGTGTGCGGCGGCGGCGGCGGCGGCGG 143
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 171 GAGGTAGTACAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 112
OY 144 GCGCTGACAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 194
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 111 CCGACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 61
```

Search completed: November 12, 2002, 16:58:57  
Job time : 36.7711 secs



QY	61	GCTTGGGCGGCAACCAATCCGGCTGAGGAGCAACACACACCTCACCGAGGAACGCTCG	120
Db	86	GCGTCGATGTCGGGCAAGCC-----CATCAAGTCGTCGACGCTGACCCAGAAAGACGCTCG	139
QY	121	TGCGCACCATCGAGTACCTGGTGGCGCTGGACGAGGAGGCCAGACCAACGATGAGACGCCGCCG	180
Db	140	TGGCCACCAATCCAAATATTTGGTTCGCTTGCGACGAGGAGTGACACACGATGACCTTCGCG	199
QY	181	GCGGCGCTCGAGGTCGCGGTCGATGTCGAGACAT	214
Db	200	GCGGCGCTCGAGGTCGCGGTCGAGAAACGACACAT	233

```

1      RESULT 2
2      PCT-US95-06790-1
3      : Sequence 1, Application PC/TUS9506790
4      : GENERAL INFORMATION:
5      :   APPLICANT: Mayo Foundation for Medical Education and Research
6      :   APPLICANT: and Hoffmann-La Roche Inc.
7      :   TITLE OF INVENTION: Detection of a Genetic Locus Encoding
8      :   TITLE OF INVENTION: Resistance to Rifampin
9      :   NUMBER OF SEQUENCES: 15
10     : CORRESPONDENCE ADDRESS:
11     :   ADDRESSEE: Schweegman, Lundberg & Woessner
12     :   STREET: 3500 IDS Center
13     :   CITY: Minneapolis
14     :   STATE: MN
15     :   COUNTRY: USA
16     :   ZIP: 55402
17     :   COMPUTER READABLE FORM:
18     :     MEDIUM TYPE: Floppy disk
19     :     COMPUTER: IBM PC compatible
20     :     OPERATING SYSTEM: PC-DOS/MS-DOS
21     :     SOFTWARE: PatentIn Release #1.0, Version #1.25
22     :   CURRENT APPLICATION DATA:
23     :     APPLICATION NUMBER: PCT/US95/06790
24     :     FILING DATE: 26-MAY-1995
25     :   CLASSIFICATION:
26     :     ATTORNEY/AGENT INFORMATION:
27     :       NAME: Raasch, Kevin W.
28     :       REGISTRATION NUMBER: 35,651
29     :       REFERENCE/DOCKET NUMBER: 150,105WO1
30     :       TELECOMMUNICATION INFORMATION:
31     :         TELEPHONE: 612-339-0331
32     :         TELEFAX: 612-339-3061
33     :   INFORMATION FOR SEQ ID NO: 1:
34     :     SEQUENCE CHARACTERISTICS:
35     :       LENGTH: 970 base pairs
36     :       TYPE: nucleic acid
37     :       STRANDEDNESS: single
38     :       TOPOLOGY: Linear
39     :     MOLECULE TYPE: DNA
40     :   CTT-US95-06790-1

```

Query Match	60.68;	Score 129.6;	DB 5;	Length 970;
Best Local Similarity	79.08;	Pred. No. 7.2e-21;		
Matches 169; Conservative	0;	Mismatches 39;	Indels 6;	Gaps 1

OY	1	TCAAGGGAAGGCGCTACGATCTGGGCGCCGCGGGCTCGCTCAAGGGTGAACAAGAAAGCTGG	60
Db	26	TCAAGGGAAGGCGCTACGATCTGGGCGCCGCGGGCTCGCTCAAGGGTGAACAAGAAAGCTGG	85
OY	61	GCTTGGGCGGCGACCAATCCGGCTCAGGTGACCAACCAACCTTCACCGAGGAAGACGTGC	120
Db	86	GCGCTCATGTCTGGCGAAGCC-----CATCAAGTGTGTGAAGCTGACCGAAGAAGACGTGC	139
OY	121	TGCCCAACCATTCAGTAATCTGGTGGCGCTTGCACAGAGGCCAGACCAACGATGACCGCCCGC	180
Db	140	TGGCCACCAATTCAAATATCTGGTCTCCGCTTGTGCACAGAGGTCAGACCAACGATGACCGTCCGG	199
OY	181	GGCGGCTCAGAGTGCACGGTTGGATGTGGACACAAAT	214
Db	200	GGCGGCTCAGAGTGCACGGTTGGATGTGGACACAAAT	233

RESULT 3  
US-08-313-185-57  
; Sequence 57, Application US/08313185  
; Patent No. 5851763  
; GENERAL INFORMATION:  
; APPLICANT, Name:

```

1 APPLICANT: Cole, Stewart
2 APPLICANT: Young, Douglas
3 APPLICANT: Zhang, Ying
4 APPLICANT: Honore, Nadine
5 APPLICANT: Telenti, Amalio
6 APPLICANT: Bodmer, Thomas
7 TITLE OF INVENTION: Rapid Detection of Antibiotic Resistance
8 TITLE OF INVENTION: In Mycobacterium Tuberculosis
9 NUMBER OF SEQUENCES: 66
10 CORRESPONDENCE ADDRESSES:
11 ADDRESSSEE: Finnegan, Henderson, Farabow, Garrett &
12 ADDRESSSEE: Dunner
13 STREET: 1300 I Street, N.W.
14 CITY: Washington
15 STATE: D.C.
16 COUNTRY: USA
17 ZIP: 20005-3315
18 COMPUTER READABLE FORM:
19 MEDIUM TYPE: Floppy disk
20 COMPUTER: IBM PC compatible
21 OPERATING SYSTEM: PC-DOS/MS-DOS
22 SOFTWARE: Patent In Release #1.0, Version #1.25
23 CURRENT APPLICATION DATA:
24 APPLICATION NUMBER: US/08/313,185
25 FILING DATE: 12-OCT-1994
26 CLASSIFICATION: 435
27 ATTORNEY/AGENT INFORMATION:
28 NAME: Meyers, Kenneth J.
29 REGISTRATION NUMBER: 25,146
30 REFERENCE/DOCKET NUMBER: 02356, 0068-00000
31 TELECOMMUNICATION INFORMATION:
32 TELEPHONE: (202) 408-4000
33 TELEFAX: (202) 408-4400
34 INFORMATION FOR SEQ ID NO: 57:
35 SEQUENCE CHARACTERISTICS:
36 LENGTH: 3447 base pairs
37 TYPE: nucleic acid
38 STRANDEDNESS: single
39 TOPOLOGY: linear
40 MOLECULE TYPE: DNA (genomic)
41 US-08-313-185-57

```

Query Match	50.8%;	Score 108.8;	DB 2;	Length 3447;
Best Local Similarity	72.9%;	Pred. No. 3.1e-16;		
Matches 156;	Conservative	0;	Mismatches 52;	Indels 6;
				Gaps 1;

QY	1	TCAAGGAAAGACGGCTACAGATCTGGCGCCGCGGGTGGTATCAAGAGTGAACACGAAGAGCTGG	60
Db			
	809	TCAAGGAAAGACGGCTACAGACTGGCCAGGGTGGTGTTCACAAAGGTCAACAGAAAGCTCG	868
QY			
	61	GCCGTGGCGGCAACCAATCCGGCTCAGGTGACCCACACACCTCTCACACGAGGAAAGCTCG	120
Db			
	869	GGTTGGACGCCCGGTGAGTTGA-----TCACGTGCTCACGCTGACCGAAGAGATGTGC	922
QY			
	121	TGCCCAACCATTCAGTACTGTGGTGGCTGCGACGAGGGCCAGACCAACGATGACCGCCCCG	180
Db			
	923	TGCCCAACCATTAAGTACTGGTTGTGTGCATATAGGGTCAGTGCACAAATGACTGTCCAG	982
QY			
	181	GGCGCTCAGAGTCCCGGTGGATGTGGACACAT	214
Db			
	983	GTGGGTTACAACTGCACATGGAACCTATGACATAT	1016

RESULT 4  
US-09-082-614A-57

Sequence 57, Application US/09082614A  
Patent No. 6124098  
GENERAL INFORMATION:  
APPLICANT: Heym, Beate  
APPLICANT: Cole, Stewart  
APPLICANT: Young, Douglas  
APPLICANT: Zhang, Ying  
APPLICANT: Honore, Nadine  
APPLICANT: Telenti, Amalio  
APPLICANT: Bodmer, Thomas  
TITLE OF INVENTION: Rapid Detection of Antibiotic Resistance  
TITLE OF INVENTION: In Mycobacterium Tuberculosis  
NUMBER OF SEQUENCES: 66  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &  
ADDRESSEE: Dunner  
STREET: 1300 I Street, N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20005-3315  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/082,614A  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/313,185  
FILING DATE: 12-OCT-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Meyers, Kenneth J.  
REGISTRATION NUMBER: 25,146  
REFERENCE/DOCKET NUMBER: 02356.0068-00000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 408-4000  
TELEFAX: (202) 408-4400  
INFORMATION FOR SEQ ID NO: 57:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3447 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-09-082-614A-57

Query Match 50.8%; Score 108.8; DB 3; Length 3447;  
Best Local Similarity 72.9%; Pred. No. 3.1e-16;  
Matches 156; Conservative 0; Mismatches 52; Indels 6; Gaps 1;

QY 1 TCAGAGAGAGCGCTACGATCTGGCCCGCGTGGTGGTGTAGCAAGGTGAACAAGAGCTGG 60  
DB 809 TCAAGAGAGAAAGCTACGACCTGGCCAGAGTGTGCTTACAGGTCAACAAGAGCTGC 868  
QY 61 GCCTGGCGCGGACCAATCCGGCTCAAGTGACACACCACTTACCGAGAGACGCTGC 120  
DB 869 GGTTCACACCGCGGTAGTTGA-----TCACGTCGCTCCACGCTGACGAGAGAGATGTGC 922  
QY 121 TGCACACCATCATGATACCTGTGCTGCTGCACAGAGAGGCGCAGACAGCATGACCGCCCGC 180  
DB 923 TCGCCACCATAGTACTGCTGGTGTGCTGTCATGAGAGGTCACTGCAATGACTGTCCAG 982  
QY 181 GGGGGGTGAGGTGCGGTGATGTGACGACAT 214  
DB 983 GTGGGTAGAAAGTCCAGTGAAGACTGACGATAT 1016

RESULT 5  
5352575-8  
Patent No. 5352575

APPLICANT: PETROVSKIS, ERIK A.; POST, LEONARD E.; TIMMINS, JAMES G.  
TITLE OF INVENTION: PSEUDORABIES VIRUS PROTEIN  
NUMBER OF SEQUENCES: 12  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/513,282  
FILING DATE: 20-APR-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 100,817  
FILING DATE: 29-JUN-1987  
APPLICATION NUMBER: 886,260  
FILING DATE: 16-JUL-1986  
APPLICATION NUMBER: 784,787  
FILING DATE: 04-OCT-1985  
APPLICATION NUMBER: 801,799  
FILING DATE: 26-NOV-1985  
APPLICATION NUMBER: 844,113  
FILING DATE: 26-MAR-1986  
SEQ ID NO: 8  
LENGTH: 1734  
5352575-8

Query Match 19.7%; Score 42.2; DB 6; Length 1734;  
Best Local Similarity 54.1%; Pred. No. 0.18;  
Matches 86; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 56 GCTGGCCCTGGCGCGCACCAATCCGGCTCAGTGACACACCACTCAGGAGAGAGA 115  
DB 1065 GATGACCCACAAAGCGCCACGCTCCGCACTGGACTGACAGCTGTGTCACCGCGCGGA 1124  
QY 116 CGTGTGCCCAATGATGATCTGTGCGCTGCACGAGGGCCAGACCATGATGACGC 175  
DB 1125 GTACGTACCGTATCAAGAGCTGACGGCCCGCGCGCGGAGACCCCGTGGGG 1184  
QY 176 CCCGCGCGCGTGTGAGTGGCGGTGATGTGACGACAT 214  
DB 1185 CCCGCGCGCGGAGAGACGCGATCTACGTGACGCGCT 1223

RESULT 6  
US-09-335-409-1  
Sequence 1, Application US/09335409  
Patent No. 6121029  
GENERAL INFORMATION:  
APPLICANT: Schupp, Thomas  
APPLICANT: Molnar, James  
APPLICANT: Zirkle, Ross  
APPLICANT: Cyr, Devon  
APPLICANT: Goerlach, Joern  
TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES  
CURRENT APPLICATION NUMBER: US/09/335,409  
CURRENT FILING DATE: 1999-06-17  
NUMBER OF SEQ ID NOS: 30  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 1  
LENGTH: 68750  
TYPE: DNA  
ORGANISM: Sorangium cellulosum  
US-09-335-409-1

Query Match 17.9%; Score 38.4; DB 3; Length 68750;  
Best Local Similarity 52.5%; Pred. No. 1.5;  
Matches 84; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

QY 19 ATCTGGCCCGCGTGGTGTGTAAGGTGAACAAGAGCTGGCGCTGGCGCACCATC 78  
DB 21325 ATCTGGCGAGCGCGCGCGCGCACCGGAGCGGAGACCTGTCTCCGCGCTGGCGCG 21384  
QY 79 CGGCTCAGGTGACCAACCAACCACTCACAGGAGAGACGTGTCGCCACCATGAGTACC 138  
DB 21385 CCGAGCGAGCGCGCGGAGCGGCTCTCTGAGCGCGCTCTCGCGGAGCATCTCGACAG 21444

; CURRENT APPLICATION NUMBER: US/09/568,486  
 ; CURRENT FILING DATE: 2000-05-10



;; PRIOR APPLICATION NUMBER: 09/335,409  
;; PRIOR FILING DATE: 1999-06-17  
;; NUMBER OF SEQ ID NOS: 30  
;; SOFTWARE: Patent In Ver. 2.0  
;; SEQ ID NO 1  
;; LENGTH: 68750  
;; TYPE: DNA  
;; ORGANISM: Sorangium cellulosum  
US-09-568-486-1

Query Match 17.9% Score 38.4; DB 4; Length 68750;  
Best Local Similarity 52.5%; Pred. No. 1.5;  
Matches 84; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

OY 19 ATCTGCCCCGCTGGGTGCTACAGGTGAACAAGAGCTGGGCGGCGACCAATC 78  
DB 21325 ATCCGCGAGCGCGCGGCGGCGGAGCGGAGCGGAGCTGCTCGCGCGCATCTCGCAGG 21384  
OY 79 CGGCTCAGTACCAACCAACCTTACCGAGAGAGCTGTCGCCACCATCAGATACC 138  
DB 21385 CCGAGCCGAGCGCGCGGCGGCTCTCTCGAGCGGCTCTCGCGCGCATCTCGCAGG 21444  
OY 139 TGGTGGCGCTGCAGAGGCGGCGAGCAAGATCGATGACCGCC 178  
DB 21445 TGCTGGCGCTCCCGAGGCGCAAGATCGAGTGCAGCCCC 21484

RESULT 11  
US-09-568-472-1  
; Sequence 1, Application US/09568472  
; Patent No. 6358719  
; GENERAL INFORMATION:  
; APPLICANT: Schupp, Thomas  
; APPLICANT: Ligon, James  
; APPLICANT: Molnar, Istvan  
; APPLICANT: Zirkle, Ross  
; APPLICANT: Cyr, Devon  
; APPLICANT: Goerlach, Joern  
; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES  
; FILE REFERENCE: 4-30582A  
; CURRENT APPLICATION NUMBER: US/09/568,472  
; CURRENT FILING DATE: 2000-05-10  
; PRIOR APPLICATION NUMBER: 09/335,409  
; PRIOR FILING DATE: 1999-06-17  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 68750  
; TYPE: DNA  
; ORGANISM: Sorangium cellulosum  
US-09-568-472-1

Query Match 17.9% Score 38.4; DB 4; Length 68750;  
Best Local Similarity 52.5%; Pred. No. 1.5;  
Matches 84; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

OY 19 ATCTGCCCCGCTGGGTGCTACAGGTGAACAAGAGCTGGGCGGCGACCAATC 78  
DB 21325 ATCCGCGAGCGCGCGGCGGCGGAGCGGAGCGGAGCTGCTCGCGCGCATCTCGCAGG 21384  
OY 79 CGGCTCAGTACCAACCAACCTTACCGAGAGAGCTGTCGCCACCATCAGATACC 138  
DB 21385 CCGAGCCGAGCGCGCGGCGGCTCTCTCGAGCGGCTCTCGCGCGCATCTCGCAGG 21444  
OY 139 TGGTGGCGCTGCAGAGGCGGCGAGCAAGATCGATGACCGCC 178  
DB 21445 TGCTGGCGCTCCCGAGGCGCAAGATCGAGTGCAGCCCC 21484

RESULT 12  
US-09-567-899-1  
; Sequence 1, Application US/09567899  
; Patent No. 6383787

;; GENERAL INFORMATION:  
;; APPLICANT: Schupp, Thomas  
;; APPLICANT: Ligon, James  
;; APPLICANT: Molnar, Istvan  
;; APPLICANT: Zirkle, Ross  
;; APPLICANT: Cyr, Devon  
;; APPLICANT: Goerlach, Joern  
;; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES  
;; FILE REFERENCE: 4-30582A  
;; CURRENT APPLICATION NUMBER: US/09/567,899  
;; CURRENT FILING DATE: 2000-05-10  
;; PRIOR APPLICATION NUMBER: 09/335,409  
;; PRIOR FILING DATE: 1999-06-17  
;; NUMBER OF SEQ ID NOS: 30  
;; SOFTWARE: Patent In Ver. 2.0  
;; SEQ ID NO 1  
;; LENGTH: 68750  
;; TYPE: DNA  
;; ORGANISM: Sorangium cellulosum  
US-09-567-899-1

Query Match 17.9% Score 38.4; DB 4; Length 68750;  
Best Local Similarity 52.5%; Pred. No. 1.5;  
Matches 84; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

OY 19 ATCTGCCCCGCTGGGTGCTACAGGTGAACAAGAGCTGGGCGGCGACCAATC 78  
DB 21325 ATCCGCGAGCGCGCGGCGGCGGAGCGGAGCGGAGCTGCTCGCGCGCATCTCGCAGG 21384  
OY 79 CGGCTCAGTACCAACCAACCTTACCGAGAGAGCTGTCGCCACCATCAGATACC 138  
DB 21385 CCGAGCCGAGCGCGCGGCGGCTCTCTCGAGCGGCTCTCGCGCGCATCTCGCAGG 21444  
OY 139 TGGTGGCGCTGCAGAGGCGGCGAGCAAGATCGATGACCGCC 178  
DB 21445 TGCTGGCGCTCCCGAGGCGCAAGATCGAGTGCAGCCCC 21484

RESULT 13  
US-08-881-784-8  
; Sequence 8, Application US/08881784  
; Patent No. 6083731  
; GENERAL INFORMATION:  
; APPLICANT: Croteau, Rodney B.  
; APPLICANT: Lupien, Shari L.  
; APPLICANT: Karp, Frank  
; TITLE OF INVENTION: RECOMBINANT MATERIALS AND METHODS FOR  
; NUMBER OF SEQUENCES: 58  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Christensen, O'Connor, Johnson and Kindness  
; STREET: 1420 Fifth Avenue, Suite 2800  
; CITY: Seattle  
; STATE: WA  
; COUNTRY: USA  
; ZIP: 98101  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/881,784  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Shelton, Dennis K.  
; REGISTRATION NUMBER: 26,997  
; REFERENCE/DOCKET NUMBER: WSUR19777  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 224-0718  
; TELEFAX: (206) 224-0779

INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1665 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
ORIGINAL SOURCE:  
ORGANISM: Mentha x piperita  
IMMEDIATE SOURCE:  
CLONE: PM17  
US-08-881-784-8

Query Match  
Best Local Similarity 51.8%; Score 38; DB 3; Length 1665;  
Matches 86; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

OY 23 GCGCCGCGTGGTCAAGGTGAACAGAGCTGGGCGGCGCAATCCGGC 82  
DB 270 GTCCCGCGAGGCGAAGAGCGGATGAGCTGTGAGCCCGCGGCGGCGTT 329  
OY 83 TCAGGTGACACACACCTCTACCGAGAGAGCTGCGCCACCTCGAGTACTGTT 142  
DB 330 CGAGAGCATCGGAGCAGATCATGTGTACGACACGACGACATCTTCAAGCCCTA 389  
OY 143 GCGCTGACAGAGGCGACACGATGACCGCCCGCGGCGGCTC 188  
DB 390 CAGCGTGCACGTGGCGCCAGATGCGAAGATCTCGCTCCGAGCTC 435

## RESULT 14

US-09-292-768-3  
Sequence 3, Application US/09292768  
Patent No. 6194185  
GENERAL INFORMATION:  
APPLICANT: Croteau, Rodney B  
APPLICANT: Lupien, Shari L  
APPLICANT: Karp, Frank  
TITLE OF INVENTION: LIMONENE HYDROXYLASES  
TITLE OF INVENTION: RECOMBINANT MATERIALS AND METHODS FOR THE PRODUCTION OF  
FILE REFERENCE: wsur13463  
CURRENT APPLICATION NUMBER: US/09/292,768  
CURRENT FILING DATE: 1999-04-14  
EARLIER APPLICATION NUMBER: 08/881,784  
EARLIER FILING DATE: 1997-06-24  
NUMBER OF SEQ ID NOS: 70  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 3  
LENGTH: 1665  
TYPE: DNA  
ORGANISM: Mentha piperita  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (19)..(1518)  
US-09-292-768-3

Query Match  
Best Local Similarity 17.8%; Score 38; DB 4; Length 1665;  
Matches 86; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

OY 23 GCGCCGCGTGGTCAAGGTGAACAGAGCTGGGCGGCGCAATCCGGC 82  
DB 270 GTCCCGCGAGGCGAAGAGCGGATGAGCTGTGAGCCCGCGGCGGCGTT 329  
OY 83 TCAGGTGACACACACCTCTACCGAGAGAGCTGCGCCACCTCGAGTACTGTT 142  
DB 330 CGAGAGCATCGGAGCAGATCATGTGTACGACACGACGACATCTTCAAGCCCTA 389  
OY 143 GCGCTGACAGAGGCGACACGATGACCGCCCGCGGCGGCTC 188  
DB 390 CAGCGTGCACGTGGCGCCAGATGCGAAGATCTCGCTCCGAGCTC 435

RESULT 15  
US-09-292-768-67  
Sequence 67, Application US/09292768  
Patent No. 6194185  
GENERAL INFORMATION:  
APPLICANT: Croteau, Rodney B  
APPLICANT: Lupien, Shari L  
APPLICANT: Karp, Frank  
TITLE OF INVENTION: LIMONENE HYDROXYLASES  
TITLE OF INVENTION: RECOMBINANT MATERIALS AND METHODS FOR THE PRODUCTION OF  
FILE REFERENCE: wsur13463  
CURRENT APPLICATION NUMBER: US/09/292,768  
CURRENT FILING DATE: 1999-04-14  
EARLIER APPLICATION NUMBER: 08/881,784  
EARLIER FILING DATE: 1997-06-24  
NUMBER OF SEQ ID NOS: 70  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 67  
LENGTH: 1665  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence:  
OTHER INFORMATION: Computer-generated nucleic acid sequence encoding  
OTHER INFORMATION: Limonene-3-hydroxylase  
FEATURE:  
NAME/KEY: misc-feature  
LOCATION: (1)..(1665)  
OTHER INFORMATION: Computer-generated nucleic acid sequence encoding  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (19)..(1518)  
US-09-292-768-67

Query Match  
Best Local Similarity 17.8%; Score 38; DB 4; Length 1665;  
Matches 86; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

OY 23 GCGCCGCGTGGTCAAGGTGAACAGAGCTGGGCGGCGCAATCCGGC 82  
DB 270 GTCCCGCGAGGCGAAGAGCGGATGAGCTGTGAGCCCGCGGCGGCGTT 329  
OY 83 TCAGGTGACACACACCTCTACCGAGAGAGCTGCGCCACCTCGAGTACTGTT 142  
DB 330 CGAGAGCATCGGAGCAGATCATGTGTACGACACGACGACATCTTCAAGCCCTA 389  
OY 143 GCGCTGACAGAGGCGACACGATGACCGCCCGCGGCGGCTC 188  
DB 390 CAGCGTGCACGTGGCGCCAGATGCGAAGATCTCGCTCCGAGCTC 435

Search completed: November 12, 2002, 21:03:28  
Job time: 236.698 secs

GenCore version 5.1.3  
Copyright (c) 1991 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 12, 2002, 15:44:35 : Search time 140.151 Seconds

(without alignments)  
3438.621 Million cell updates/sec

Title: US-09-697-123b-21

Perfect score: 214

Sequence: 1 tcaaggagaagcgctacgat.....ccggtggtatgtgacacacat 214

Scoring table:

IDENTITY\_NUC  
Gapop 10.0 , Gapept 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: N\_Geneseq\_101002.\*  
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:\*  
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:\*  
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:\*  
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:\*  
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:\*  
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:\*  
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:\*  
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:\*  
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:\*  
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:\*  
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:\*  
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:\*  
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:\*  
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:\*  
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:\*  
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:\*  
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:\*  
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:\*  
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:\*  
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:\*  
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:\*  
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:\*  
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	214	100.0	214	22	Myobacterium absc
2	186.8	87.3	214	22	Myobacterium chel
3	153.6	71.8	208	22	Myobacterium fort
4	143.6	68.0	208	22	Myobacterium aviu
5	145.2	67.9	214	22	Myobacterium maln
6	142.4	66.5	208	22	Myobacterium gord
7	142.4	66.5	208	22	Myobacterium kans
8	142.4	66.5	208	22	Myobacterium ulce
9	142	66.4	211	22	Myobacterium flav

10	140.8	65.8	208	22	AA505211
11	140.8	65.8	208	22	AA505217
12	137.6	64.3	208	22	AA505202
13	136	63.6	208	22	AA505203
14	136	63.6	208	22	AA505224
15	129.6	60.6	208	22	AA505205
16	129.6	60.6	208	22	AA505206
17	129.6	60.6	208	22	AA505216
18	129.6	60.6	208	22	AA505216
19	129.6	60.6	3519	17	AA709676
20	129.6	60.6	3519	22	AAH02079
21	129.6	60.6	3853	22	AAH02079
22	129.6	60.6	3853	21	AAH74651
23	128	59.8	208	22	AAH89994
24	128	59.8	208	22	AA505222
25	121.6	56.8	205	22	AA505220
26	120.8	56.4	207	22	AA505122
27	119.2	55.7	207	22	AA505122
28	108.8	50.4	3447	14	AA505132
29	107.8	50.4	223	22	AA505213
30	103	48.1	223	22	AA505209
31	80.4	37.6	3495	22	AAH65512
32	80.4	37.6	3495	22	AAH65512
33	55.2	25.8	27426	23	AAH68525
34	43.8	20.5	1734	21	AAH40350
35	43.8	20.5	5895	23	ABH87076
36	43.2	20.2	1844	12	AA012000
37	42.8	20.0	77536	21	AAH4651
38	42.2	19.7	1734	8	AAH70995
39	42.2	19.7	1734	15	AAH73490
40	42.2	19.7	1734	22	AAH09832
41	42.2	19.7	1734	22	AAH09827
42	41.4	19.3	547	21	AAH03907
43	40.4	18.9	109519	22	AAH08693
44	40.2	18.8	537	21	AAH56489
45	39.6	18.5	2853	23	ABH12631

#### ALIGNMENTS

RESULT 1  
AA505221  
ID AA505221 standard: DNA: 214 BP.  
AC AA505221:  
DT 07-SEP-2001 (first entry)  
XX  
XX Myobacterium abscessus rpoB gene fragment.  
DE  
XX Non-tuberculous mycobacteria: rpoB gene fragment; NTM; HIV; PRA; RFLP;  
KM PCR-restriction fragment length polymorphism analysis; ds.  
OS  
XX Myobacterium abscessus.  
XX  
XX W0200131061-A1.  
XX  
XX 03-MAY-2001.  
XX  
XX 27-OCT-2000: 2000MO-KR01223.  
XX  
XX 27-OCT-1999: 99KR-0046795.  
XX  
XX (ERDM-) ERDM BIOTECH CO LTD.  
XX  
XX Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;  
XX WPI: 2001-300520/31.  
XX  
XX New DNA fragments from the rpoB gene of mycobacteria, useful for  
XX diagnosis and identification of many mycobacterial species by  
XX restriction fragment length polymorphism -  
XX PT  
XX PT

Myobacterium mari  
Myobacterium cela  
Myobacterium gord  
Myobacterium gord  
Myobacterium xeno  
Myobacterium tube  
Myobacterium teir  
Myobacterium bov1  
Myobacterium tube  
Myobacterium tube  
Myobacterium tube  
Myobacterium afri  
Myobacterium haem  
Myobacterium intr  
Myobacterium szul  
Myobacterium gord  
M. leprae rpoB gene  
Myobacterium gast  
Myobacterium scro  
C glutamicum codin  
Propionibacterium  
Swine pseudorabies  
Pseudorabies virus  
Maize nitrate redu  
Nucleotide sequenc  
Sequence encoding  
DNA encoding gi of  
Pseudorabies virus  
PRV glycoprotein 9  
Human secreted pro  
Microsporidia DNA  
Eucalyptus grandis  
Drosophila melanog

XX Claim 1; Page 46; 50pp; English.

PS The present sequence for Mycobacterium abscessus rpoB gene  
XX fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from  
CC various Mycobacterial species. These rpoB gene fragments can be used  
CC in the diagnosis and identification of Mycobacterium species using a  
CC novel PCR-restriction fragment length polymorphism analysis (PRA)  
CC method. The method comprises obtaining a restriction fragment length  
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,  
CC amplifying and digesting the DNA fragment from the microorganism to  
CC be identified and comparing the RFLP patterns from the known rpoB gene  
CC fragments with the unidentified fragment. The rpoB gene fragments  
CC are useful to identify a wide range of Mycobacterium species, e.g. for  
CC diagnosis or to obtain epidemiological and pathogenesis information for  
CC selection of appropriate therapies, including M. tuberculosis, M. leprae  
CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected  
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene  
CC fragments is rapid, precise, simple and cost effective (only 1 PCR  
CC required), and can differentiate between many species in a single  
CC experiment, including those difficult to distinguish by usual biochemical  
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for  
CC detecting specific Mycobacterial species.

SO Sequence 214 BP; 45 A; 69 C; 72 G; 28 T; 0 other;

Query Match 100.0%; Score 214; DB 22; Length 214;  
Best Local Similarity 100.0%; Pred. No. 5,4e-38;  
Matches 214; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TCAGAGAAAGCGCTACGATCTGCGCCGCTGGCTGTACAGGTGAACAAAGCTGG 60  
DB 1 TCAGAGAAAGCGCTACGATCTGCGCCGCTGGCTGTACAGGTGAACAAAGCTGG 60  
OY 61 GCCTGGGCGGCACCAATCGCGCTCAGTGTGACACACACCTCAACGAGAGACCTGC 120  
DB 61 GCCTGGGCGGCACCAATCGCGCTCAGTGTGACACACACCTCAACGAGAGACCTGC 120  
OY 121 TCGCCACCATGAGTACCTGTGCGCCTGTGACAGAGGGCCAGACCATGATGACCGCCCG 180  
DB 121 TCGCCACCATGAGTACCTGTGCGCCTGTGACAGAGGGCCAGACCATGATGACCGCCCG 180  
OY 181 GCGGCTCGAGGTGCGGCTGGATGTGAGCGACAT 214  
DB 181 GCGGCTCGAGGTGCGGCTGGATGTGAGCGACAT 214

## RESULT 2

AAS05207 standard; DNA; 214 BP.

AC AAS05207;

DT 07-SEP-2001 (first entry)

XX Mycobacterium chelonae rpoB gene fragment.

XX Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;

KW PCR-restriction fragment length polymorphism analysis; ds.

OS Mycobacterium chelonae.

PN WO200131061-A1.

PD 03-MAY-2001.

PF 27-OCT-2000; 2000WO-KR01223.

PR 27-OCT-1999; 99KR-0046795.

PA (ERUM-) ERUME BIOTECH CO LTD.

XX Lee H, Park YK, Bal G, Kim S, Cho S, Kim Y, Park HJ;

XX RFL; 2001-300520/31.

PT New DNA fragments from the rpoB gene of mycobacteria, useful for  
XX diagnosis and identification of many mycobacterial species by  
XX restriction fragment length polymorphism

PS Claim 1; Page 42; 50pp; English.

XX The present sequence for Mycobacterium chelonae rpoB gene  
CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from  
CC various Mycobacterial species. These rpoB gene fragments can be used  
CC in the diagnosis and identification of Mycobacterium species using a  
CC novel PCR-restriction fragment length polymorphism analysis (PRA)  
CC method. The method comprises obtaining a restriction fragment length  
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,  
CC amplifying and digesting the DNA fragment from the microorganism to  
CC be identified and comparing the RFLP patterns from the known rpoB gene  
CC fragments with the unidentified fragment. The rpoB gene fragments  
CC are useful to identify a wide range of Mycobacterium species, e.g. for  
CC diagnosis or to obtain epidemiological and pathogenesis information for  
CC selection of appropriate therapies, including M. tuberculosis, M. leprae  
CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected  
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene  
CC fragments is rapid, precise, simple and cost effective (only 1 PCR  
CC required), and can differentiate between many species in a single  
CC experiment, including those difficult to distinguish by usual biochemical  
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for  
CC detecting specific Mycobacterial species.

SO Sequence 214 BP; 41 A; 72 C; 72 G; 29 T; 0 other;

Query Match 87.3%; Score 186.8; DB 22; Length 214;  
Best Local Similarity 92.1%; Pred. No. 4,6e-32;  
Matches 197; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

OY 1 TCAGAGAAAGCGCTACGATCTGCGCCGCTGGCTGTACAGGTGAACAAAGCTGG 60  
DB 1 TCAGAGAAAGCGCTACGATCTGCGCCGCTGGCTGTACAGGTGAACAAAGCTGG 60  
OY 61 GCCTGGGCGGCACCAATCGCGCTCAGTGTGACACACCTCAACGAGAGACCTGC 120  
DB 61 GCCTGGGCGGCACCAATCGCGCTCAGTGTGACACACCTCAACGAGAGACCTGC 120  
OY 121 TCGCCACCATGAGTACCTGTGCGCCTGTGACAGAGGGCCAGACCATGATGACCGCCCG 180  
DB 121 TCGCCACCATGAGTACCTGTGCGCCTGTGACAGAGGGCCAGACCATGATGACCGCCCG 180  
OY 181 GCGGCTCGAGGTGCGGCTGGATGTGAGCGACAT 214  
DB 181 GCGGCTCGAGGTGCGGCTGGATGTGAGCGACAT 214

## RESULT 3

AAS05219 standard; DNA; 208 BP.

AC AAS05219;

DT 07-SEP-2001 (first entry)

XX Mycobacterium fortuitum rpoB gene fragment.

XX Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;

KW PCR-restriction fragment length polymorphism analysis; ds.

OS Mycobacterium fortuitum.

PN WO200131061-A1.

PD 03-MAY-2001.

PF 27-OCT-2000; 2000WO-KR01223.

```
XX 27-OCT-1999; 99KR-0046795.
PR (ERUM-) ERUME BIOTECH CO LTD.
XX
XX Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;
XX WPI; 2001-300520/31.
XX
PT New DNA fragments from the rpoB gene of mycobacteria, useful for
PT diagnosis and identification of many mycobacterial species by
PT restriction fragment length polymorphism
XX
XX Claim 1: Page 46; 50pp; English.
XX
XX The present sequence for Mycobacterium fortuitum rpoB gene
XX fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
XX various Mycobacterial species. These rpoB gene fragments can be used
XX in the diagnosis and identification of Mycobacterial species using a
XX novel PCR-restriction fragment length polymorphism analysis (PRA)
XX method. The method comprises obtaining a restriction fragment length
XX polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
XX amplifying and digesting the DNA fragment from the microorganism to
XX be identified and comparing the RFLP patterns from the known rpoB gene
XX fragments with the unidentified fragment. The rpoB gene fragments
XX are useful to identify a wide range of Mycobacterium species, e.g. for
XX diagnosis or to obtain epidemiological and pathogenesis information for
XX selection of appropriate therapies, including M. tuberculosis, M. leprae
XX and non-tuberculous mycobacteria (NTM) encountered in subjects infected
XX with human immunodeficiency virus (HIV). Analysis of the rpoB gene
XX fragments is rapid, precise, simple and cost effective (only 1 PCR
XX experiment), and can differentiate between many species in a single
XX tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
XX detecting specific Mycobacterial species.
XX
SQ Sequence 208 BP; 43 A; 69 C; 69 G; 27 T; 0 other:
Query Match 71.8%; Score 153.6; DB 22; Length 208;
Best Local Similarity 86.0%; Pred. No. 7.8e-25;
Matches 184; Conservative 0; Mismatches 24; Indels 6; Gaps 1;
QY 1 TCAAGAGAACGCGTACGATCTGCGCGGTCGCTACAGCGTGAACAGAACTGG 60
DB 1 TCAAGAGAACGCGTACGATCTGCGCGGTCGCTACAGCGTGAACAGAACTGG 60
QY 61 GCTGGGCGGACCAATCGCGCTCAGGTGACCAACCAACCTCACCAGAGAAAGCTCG 120
DB 61 GCTGGGCGGACCAATCGCGCTCAGGTGACCAACCAACCTCACCAGAGAAAGCTCG 114
QY 121 TCGGCACCATCGATGACTGCTGCGCTGACAGAGGCGCCACAGACGATGACCGCCCG 180
DB 115 TCGGCACCATCGATGACTGCTGCGCTGACAGAGGCGCCACAGACGATGACCGCCCG 174
QY 181 GCGGCGTGCAGGTGCGCGGTGATGTGAGACAT 214
DB 175 GCGGCGTGCAGGTGCGCGGTGATGTGAGACAT 208
RESULT 4
AAS05215
ID AAS05215 standard; DNA: 208 BP.
XX
XX AAS05215;
XX
XX 07-SEP-2001 (first entry)
XX
XX Mycobacterium avium rpoB gene fragment.
XX
XX Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;
XX PCR-restriction fragment length polymorphism analysis; DS.
XX
XX Mycobacterium avium.
```

```
XX
XX MO200131061-A1.
XX
XX 03-MAY-2001.
XX
XX 27-OCT-2000; 2000WO-KR01223.
XX
XX 27-OCT-1999; 99KR-0046795.
XX
XX (ERUM-) ERUME BIOTECH CO LTD.
XX
XX Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;
XX WPI; 2001-300520/31.
XX
XX New DNA fragments from the rpoB gene of mycobacteria, useful for
XX diagnosis and identification of many mycobacterial species by
XX restriction fragment length polymorphism
XX
XX Claim 1: Page 44; 50pp; English.
XX
XX The present sequence for Mycobacterium avium rpoB gene
XX fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
XX various Mycobacterial species. These rpoB gene fragments can be used
XX in the diagnosis and identification of Mycobacterial species using a
XX novel PCR-restriction fragment length polymorphism analysis (PRA)
XX method. The method comprises obtaining a restriction fragment length
XX polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
XX amplifying and digesting the DNA fragment from the microorganism to
XX be identified and comparing the RFLP patterns from the known rpoB gene
XX fragments with the unidentified fragment. The rpoB gene fragments
XX are useful to identify a wide range of Mycobacterium species, e.g. for
XX diagnosis or to obtain epidemiological and pathogenesis information for
XX selection of appropriate therapies, including M. tuberculosis, M. leprae
XX and non-tuberculous mycobacteria (NTM) encountered in subjects infected
XX with human immunodeficiency virus (HIV). Analysis of the rpoB gene
XX fragments is rapid, precise, simple and cost effective (only 1 PCR
XX experiment), and can differentiate between many species in a single
XX tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
XX detecting specific Mycobacterial species.
XX
SQ Sequence 208 BP; 44 A; 69 C; 69 G; 26 T; 0 other:
Query Match 68.0%; Score 145.6; DB 22; Length 208;
Best Local Similarity 83.0%; Pred. No. 4.3e-23;
Matches 179; Conservative 0; Mismatches 29; Indels 6; Gaps 1;
QY 1 TCAAGAGAACGCGTACGATCTGCGCGGTCGCTACAGCGTGAACAGAACTGG 60
DB 1 TCAAGAGAACGCGTACGATCTGCGCGGTCGCTACAGCGTGAACAGAACTGG 60
QY 61 GCTGGGCGGACCAATCGCGCTCAGGTGACCAACCAACCTCACCAGAGAAAGCTCG 120
DB 61 GCTGGGCGGACCAATCGCGCTCAGGTGACCAACCAACCTCACCAGAGAAAGCTCG 114
QY 121 TCGGCACCATCGATGACTGCTGCGCTGACAGAGGCGCCACAGACGATGACCGCCCG 180
DB 115 TCGGCACCATCGATGACTGCTGCGCTGACAGAGGCGCCACAGACGATGACCGCCCG 174
QY 181 GCGGCGTGCAGGTGCGCGGTGATGTGAGACAT 214
DB 175 GCGGCGTGCAGGTGCGCGGTGATGTGAGACAT 208
RESULT 5
AAS05214
ID AAS05214 standard; DNA: 214 BP.
XX
XX AAS05214;
XX
XX 07-SEP-2001 (first entry)
XX
```

DE Mycobacterium malmoeense rpoB gene fragment.

XX Non-tuberculous mycobacteria: rpoB gene fragment; NTM; HIV; PRA; RFLP;  
KM PCR-restriction fragment length polymorphism analysis; ds.  
XX

OS Mycobacterium malmoeense.

XX W0200131061-A1.

XX 03-MAY-2001.

XX 27-OCT-2000; 2000MO-KR01223.

XX 27-OCT-1999; 99KR-0046795.

XX (ERUM-) ERMUE BIOTECH CO LTD.

XX Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;

XX WPI; 2001-300520/31.

PT New DNA fragments from the rpoB gene of mycobacteria, useful for  
PT diagnosis and identification of many mycobacterial species by  
PT restriction fragment length polymorphism  
XX  
XX Claim 1; Page 44; 50pp; English.

XX The present sequence for Mycobacterium malmoeense rpoB gene  
CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from  
CC various Mycobacterial species. These rpoB gene fragments can be used  
CC in the diagnosis and identification of Mycobacterium species using a  
CC novel PCR-restriction fragment length polymorphism analysis (PRA)  
CC method. The method comprises obtaining a restriction fragment length  
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,  
CC amplifying and digesting the DNA fragment from the microorganism to  
CC be identified and comparing the RFLP patterns from the known rpoB gene  
CC fragments with the unidentified fragment. The rpoB gene fragments  
CC are useful to identify a wide range of Mycobacterium species, e.g. for  
CC selection of appropriate epidemiological and pathogenesis information for  
CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected  
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene  
CC fragments is rapid, precise, simple and cost effective (only 1 PCR  
CC experiment), and can differentiate between many species in a single  
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for  
CC detecting specific Mycobacterial species.

XX Sequence 214 BP; 42 A; 68 C; 75 G; 29 T; 0 other;

XX Query Match 67.9%; Score 145.2; DB 22; Length 214;

XX Best Local Similarity 79.9%; Pred. No. 5.3e-23;

XX Matches 171; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

XX 1 TCAAGAGAGAGCGCTACGATCTGCGCCCGCTGGCTGACAGGTGACAAAGAGCTG 60

XX 1 TCAAGAGAGAGCGCTACGATCTGCGCCCGCTGGCTGACAGGTGACAAAGAGACTG 60

XX 61 GCGTGGGCGGACCAATCTGAGTGCAGTACACCAACACCTTACCGGAGAGAGCTG 120

XX 61 GCGTGGGCGGCGGAGTGCAGTACACCAACCAACCTTACCGGAGAGAGCTG 120

XX 121 TCGCCACCATGAGTACCTGCTGCGCTGACGAGGCGCAGACAGATGAGCCGCGG 180

XX 121 TCGCCACCATGAGTACCTGCTGCGCTGACGAGGCGCAGACAGATGAGCCGCGG 180

XX 181 GCGGCGTGAAGTCCCGGTGAGACGACGACAT 214

XX 181 GCGGCGTGAAGTCCCGGTGAGACGACGACAT 214

RESULT 6  
AAS05201

ID AAS05201 standard; DNA: 208 BP.

XX AAS05201;

XX 07-SEP-2001 (first entry)

XX Mycobacterium gordonae type I rpoB gene fragment.

XX Non-tuberculous mycobacteria: rpoB gene fragment; NTM; HIV; PRA; RFLP;  
KM PCR-restriction fragment length polymorphism analysis; ds.  
XX

OS Mycobacterium gordonae type I.

XX W0200131061-A1.

XX 03-MAY-2001.

XX 27-OCT-2000; 2000MO-KR01223.

XX 27-OCT-1999; 99KR-0046795.

XX (ERUM-) ERMUE BIOTECH CO LTD.

XX Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;

XX WPI; 2001-300520/31.

PT New DNA fragments from the rpoB gene of mycobacteria, useful for  
PT diagnosis and identification of many mycobacterial species by  
PT restriction fragment length polymorphism  
XX  
XX Claim 1; Page 40; 50pp; English.

XX The present sequence for Mycobacterium gordonae type I rpoB gene  
CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from  
CC various Mycobacterial species. These rpoB gene fragments can be used  
CC in the diagnosis and identification of Mycobacterium species using a  
CC novel PCR-restriction fragment length polymorphism analysis (PRA)  
CC method. The method comprises obtaining a restriction fragment length  
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,  
CC amplifying and digesting the DNA fragment from the microorganism to  
CC be identified and comparing the RFLP patterns from the known rpoB gene  
CC fragments with the unidentified fragment. The rpoB gene fragments  
CC are useful to identify a wide range of Mycobacterium species, e.g. for  
CC selection of appropriate epidemiological and pathogenesis information for  
CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected  
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene  
CC fragments is rapid, precise, simple and cost effective (only 1 PCR  
CC experiment), and can differentiate between many species in a single  
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for  
CC detecting specific Mycobacterial species.

XX Sequence 208 BP; 46 A; 71 C; 65 G; 26 T; 0 other;

XX Query Match 66.5%; Score 142.4; DB 22; Length 208;

XX Best Local Similarity 82.7%; Pred. No. 2.1e-22;

XX Matches 177; Conservative 0; Mismatches 31; Indels 6; Gaps 1;

XX 1 TCAAGAGAGAGCGCTACGATCTGCGCCCGCTGGCTGATCAAGGTGACAAAGAGCTG 60

XX 1 TCAAGAGAGAGCGCTACGATCTGCGCCCGCTGGCTGATCAAGGTGACAAAGAGCTG 60

XX 61 GCGTGGGCGGACCAATCTGAGTGCAGTACACCAACCAACCTTACCGGAGAGAGCTG 120

XX 61 GCGTGGGCGGCGGAGTGCAGTACACCAACCAACCTTACCGGAGAGAGCTG 120

XX 121 TCGCCACCATGAGTACCTGCTGCGCTGACGAGGCGCAGACAGATGAGCCGCGG 180

XX 121 TCGCCACCATGAGTACCTGCTGCGCTGACGAGGCGCAGACAGATGAGCCGCGG 180

XX 181 GCGGCGTGAAGTCCCGGTGAGACGACGACAT 214

XX 181 GCGGCGTGAAGTCCCGGTGAGACGACGACAT 214

Db 175 GCGGACCGAGGTCCGCTGTAGACGACGACAT 208  
||||| ||||||| ||| |||||||  
RESULT 7  
ID AAS05208 standard; DNA: 208 BP.  
XX  
XX AAS05208:  
XX  
XX 07-SEP-2001 (first entry)  
XX  
XX Mycobacterium kansasii rpoB gene fragment.  
XX  
XX Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;  
XX PCR-restriction fragment length polymorphism analysis; ds.  
XX Mycobacterium kansasii.  
XX  
XX W0200131061-A1.  
XX  
XX 03-MAY-2001.  
XX  
XX 27-OCT-2000: 2000MO-KR01223.  
XX  
XX 27-OCT-1999: 99KR-0046795.  
XX  
XX (ERUM-) ERUME BIOTECH CO LTD.  
XX  
XX Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;  
XX  
XX MPI: 2001-300520/31.  
XX  
XX New DNA fragments from the rpoB gene of mycobacteria, useful for  
XX diagnosis and identification of many mycobacterial species by  
XX restriction fragment length polymorphism -  
XX  
XX Claim 1: Page 42: 50pp: English.  
XX  
XX The present sequence for Mycobacterium kansasii rpoB gene  
XX fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from  
XX various Mycobacterial species. These rpoB gene fragments can be used  
XX in the diagnosis and identification of Mycobacterium species using a  
XX novel PCR-restriction fragment length polymorphism analysis (PRA)  
XX method. The method comprises obtaining a restriction fragment length  
XX polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,  
XX amplifying and digesting the DNA fragment from the microorganism to  
XX be identified and comparing the RFLP patterns from the known rpoB gene  
XX fragments with the unidentified fragment. The rpoB gene fragments  
XX are useful to identify a wide range of Mycobacterium species, e.g. for  
XX diagnosis or to obtain epidemiological and pathogenesis information for  
XX selection of appropriate therapies, including M. tuberculosis, M. leprae  
XX and non-tuberculous mycobacteria (NTM) encountered in subjects infected  
XX with human immunodeficiency virus (HIV). Analysis of the rpoB gene  
XX fragments is rapid, precise, simple and cost effective (only 1 PCR  
XX required), and can differentiate between many species in a single  
XX experiment, including those difficult to distinguish by usual biochemical  
XX tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for  
XX detecting specific Mycobacterial species.  
XX  
XX Sequence 208 BP: 51 A; 65 C; 65 G; 27 T; 0 other:  
SQ  
Query Match 66.5%; Score 142.4; DB 22: Length 208:  
Best Local Similarity 82.7%; Pred. No. 2.le-22;  
Matches 177: Conservative 0; Mismatches 31; Indels 6; Gaps 1;  
OY 1 TCAAGAGAGAGCGTACGATCTGCGCCGCTGGTCAAGTGAACAGAGCTGG 60  
DB 1 TCAAGAGAGAGCGTACGATCTGCGCCGCTGGTCAAGTGAACAGAGCTGG 60  
OY 61 GCGTGGGCGGACCAATCGGCTAGGTGAGTACGACCACTCCCTACCGAGGAGAGCTCG 120  
DB 61 GCGTGGGCGGACCAATCGGCTAGGTGAGTACGACCACTCCCTACCGAGGAGAGCTCG 114

OY 121 TCGCCACCATCGAGTACGTGCGCTGCGACGAGGCGCCAGACCAAGTACCGCCCG 180  
DB 115 TCGCCACCATCGAGTACGTGCGCTGCGACGAGGCGCCAGACCAAGTACCGCTGCG 174  
OY 181 GCGGCGTGGAGGTGCGCGGTGAGTGTGAGACGACAT 214  
DB 175 GCGGCGTGGAGGTGCGCGGTGAGTGTGAGACGACAT 208  
||||| ||||||| ||||||| |||||||  
RESULT 8  
ID AAS05210 standard; DNA: 208 BP.  
XX  
XX AAS05210:  
XX  
XX 07-SEP-2001 (first entry)  
XX  
XX Mycobacterium ulcerans rpoB gene fragment.  
XX  
XX Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;  
XX PCR-restriction fragment length polymorphism analysis; ds.  
XX Mycobacterium ulcerans.  
XX  
XX W0200131061-A1.  
XX  
XX 03-MAY-2001.  
XX  
XX 27-OCT-2000: 2000MO-KR01223.  
XX  
XX 27-OCT-1999: 99KR-0046795.  
XX  
XX (ERUM-) ERUME BIOTECH CO LTD.  
XX  
XX Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;  
XX  
XX MPI: 2001-300520/31.  
XX  
XX New DNA fragments from the rpoB gene of mycobacteria, useful for  
XX diagnosis and identification of many mycobacterial species by  
XX restriction fragment length polymorphism -  
XX  
XX Claim 1: Page 43: 50pp: English.  
XX  
XX The present sequence for Mycobacterium ulcerans rpoB gene  
XX fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from  
XX various Mycobacterial species. These rpoB gene fragments can be used  
XX in the diagnosis and identification of Mycobacterium species using a  
XX novel PCR-restriction fragment length polymorphism analysis (PRA)  
XX method. The method comprises obtaining a restriction fragment length  
XX polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,  
XX amplifying and digesting the DNA fragment from the microorganism to  
XX be identified and comparing the RFLP patterns from the known rpoB gene  
XX fragments with the unidentified fragment. The rpoB gene fragments  
XX are useful to identify a wide range of Mycobacterium species, e.g. for  
XX diagnosis or to obtain epidemiological and pathogenesis information for  
XX selection of appropriate therapies, including M. tuberculosis, M. leprae  
XX and non-tuberculous mycobacteria (NTM) encountered in subjects infected  
XX with human immunodeficiency virus (HIV). Analysis of the rpoB gene  
XX fragments is rapid, precise, simple and cost effective (only 1 PCR  
XX required), and can differentiate between many species in a single  
XX experiment, including those difficult to distinguish by usual biochemical  
XX tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for  
XX detecting specific Mycobacterial species.  
XX  
XX Sequence 208 BP: 45 A; 68 C; 67 G; 28 T; 0 other:  
SQ  
Query Match 66.5%; Score 142.4; DB 22: Length 208:  
Best Local Similarity 82.7%; Pred. No. 2.le-22;  
Matches 177: Conservative 0; Mismatches 31; Indels 6; Gaps 1;  
OY 1 TCAAGAGAGAGCGTACGATCTGCGCCGCTGGTCAAGTGAACAGAGCTGG 60

Sequence 211 BP; 49 A; 65 C; 70 G; 27 T; 0 other;

with human immunodeficiency virus (HIV). Analysis of the proviral



CC required), and can differentiate between many species in a single  
CC experiment, including those difficult to distinguish by usual biochemical  
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for  
CC detecting specific Mycobacterial species.

SO Sequence 208 BP; 45 A; 69 C; 68 G; 26 T; 0 other;

Query Match 65.8%; Score 140.8; DB 22; Length 208;  
Best Local Similarity 82.2%; Pred. No. 4.8e-22;

Matches 176: Conservative 0; Mismatches 32; Indels 6; Gaps 1;

0Y 1 TCAGAGAGAGCGCTACGATCTGCGCCCGTGGCTGTACAGGTGAACAAGACTGG 60  
DB 1 TCAGAGAGAGCGCTACGATCTGCGCCCGTGGCTGTACAGGTGAACAAGACTGG 60  
0Y 61 GCCTGGCGGCGACCAATCCGCTCAGTGACACCAACCCCTGACCGAGAGAGACTCG 120  
DB 61 GCCTGGAGCGCGGCGACGCC-----CATCACCACTGTGACCGAGAGAAACCTCG 114  
0Y 121 TCGCCACCATGAGTACTGTGCGCCTGCACAGAGGCGCCAGACCCAGATGACCGCCCG 180  
DB 115 TCGCCACCATGAGTACTGTGCGCCTGCACAGAGGCGCCAGACCGGATGACCGCTCCG 174  
0Y 181 GCGGCGTGCAGGTGCGGCTGATGTGAGAGCAGCAT 214  
DB 175 GCGGCGTGCAGGTGCGGCTGATGTGAGAGCAGCAT 208

RESULT 11

AAS05217  
ID AAS05217 standard; DNA; 208 BP.

AC AAS05217;

DT 07-SEP-2001 (first entry)

DE Mycobacterium celatum rpoB gene fragment.

KW Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;

KM PCR-restriction fragment length polymorphism analysis; ds.

OS Mycobacterium celatum.

PN WO200131061-A1.

PD 03-MAY-2001.

PF 27-OCT-2000; 2000WO-KR01223.

PR 27-OCT-1999; 99KR-0046795.

PA (ERUM-) ERUME BIOTECH CO LTD.

PI Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;

DR WPI; 2001-300520/31.

PT New DNA fragments from the rpoB gene of mycobacteria, useful for  
PT diagnosis and identification of many mycobacterial species by  
PT restriction fragment length polymorphism -  
XX  
XX  
PS Claim 1; Page 45; 50pp; English.

CC The present sequence for Mycobacterium celatum rpoB gene  
CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from  
CC various Mycobacterial species. These rpoB gene fragments can be used  
CC in the diagnosis and identification of Mycobacterium species using a  
CC novel PCR-restriction fragment length polymorphism analysis (PRA)  
CC method. The method comprises obtaining a restriction fragment length  
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,  
CC amplifying and digesting the DNA fragment from the microorganism to  
CC be identified and comparing the RFLP patterns from the known rpoB gene  
CC fragments with the unidentified fragment. The rpoB gene fragments

CC are useful to identify a wide range of Mycobacterium species, e.g. for  
CC diagnosis or to obtain epidemiological and pathogenesis information for  
CC selection of appropriate therapies, including M. tuberculosis, M. leprae  
CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected  
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene  
CC fragments is rapid, precise, simple and cost effective (only 1 PCR  
CC required), and can differentiate between many species in a single  
CC experiment, including those difficult to distinguish by usual biochemical  
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for  
CC detecting specific Mycobacterial species.

SO Sequence 208 BP; 48 A; 71 C; 64 G; 25 T; 0 other;

Query Match 65.8%; Score 140.8; DB 22; Length 208;  
Best Local Similarity 82.2%; Pred. No. 4.8e-22;

Matches 176: Conservative 0; Mismatches 32; Indels 6; Gaps 1;

0Y 1 TCAGAGAGAGCGCTACGATCTGCGCCCGTGGCTGTACAGGTGAACAAGACTGG 60  
DB 1 TCAGAGAGAGCGCTACGATCTGCGCCCGTGGCTGTACAGGTGAACAAGACTGG 60  
0Y 61 GCCTGGCGGCGACCAATCCGCTCAGTGACACCAACCCCTGACCGAGAGAGACTCG 120  
DB 61 GCCTG-----AACACCGCGTCCCGCATCACAGACCACTGTGACCGAGAGAGACTCG 114  
0Y 121 TCGCCACCATGAGTACTGTGCGCCTGCACAGAGGCGCCAGACAGATGACCGCCCG 180  
DB 115 TCGCCACCATGAGTACTGTGCGCCTGCACAGAGGCGCCAGACAGATGACCGCTCCG 174  
0Y 181 GCGGCGTGCAGGTGCGGCTGATGTGAGAGCAGCAT 214  
DB 175 GCGGCGTGCAGGTGCGGCTGATGTGAGAGCAGCAT 208

RESULT 12

AAS05202  
ID AAS05202 standard; DNA; 208 BP.

AC AAS05202;

DT 07-SEP-2001 (first entry)

DE Mycobacterium gordonae type II rpoB gene fragment.

KW Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;

KM PCR-restriction fragment length polymorphism analysis; ds.

OS Mycobacterium gordonae type II.

PN WO200131061-A1.

PD 03-MAY-2001.

PF 27-OCT-2000; 2000WO-KR01223.

PR 27-OCT-1999; 99KR-0046795.

PA (ERUM-) ERUME BIOTECH CO LTD.

PI Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;

DR WPI; 2001-300520/31.

PT New DNA fragments from the rpoB gene of mycobacteria, useful for  
PT diagnosis and identification of many mycobacterial species by  
PT restriction fragment length polymorphism -  
XX  
XX  
PS Claim 1; Page 40; 50pp; English.

CC The present sequence for Mycobacterium gordonae type II rpoB gene  
CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from  
CC various Mycobacterial species. These rpoB gene fragments can be used  
CC in the diagnosis and identification of Mycobacterium species using a



DR WPI: 2001-300520/31.

XX New DNA fragments from the rpoB gene of mycobacteria, useful for  
PT diagnosis and identification of many mycobacterial species by  
restriction fragment length polymorphism

PS Claim 1: Page 47: 50pp: English.

XX The present sequence for Mycobacterium xenopi rpoB gene  
CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from  
CC various Mycobacterial species. These rpoB gene fragments can be used  
CC in the diagnosis and identification of Mycobacterium species using a  
CC novel PCR-restriction fragment length polymorphism analysis (PRA)  
CC method. The method comprises obtaining a restriction fragment length  
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,  
CC amplifying and digesting the DNA fragment from the microorganism to  
CC be identified and comparing the RFLP patterns from the known rpoB gene  
CC fragments with the unidentified fragment. The rpoB gene fragments  
CC are useful to identify a wide range of Mycobacterium species, e.g. for  
CC diagnosis or to obtain epidemiological and pathogenesis information for  
CC selection of appropriate therapies, including M. tuberculosis, M. leprae  
CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected  
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene  
CC fragments is rapid, precise, simple and cost effective (only 1 PCR  
CC experiment), and can differentiate between many species in a single  
CC test. Also described are oligonucleotide probes (AAS05227-AAS05242) for  
CC detecting specific Mycobacterial species.

SQ Sequence 208 BP: 51 A; 65 C; 67 G; 25 T; 0 other;

Query Match 63.6%; Score 136; DB 22; Length 208;  
Best Local Similarity 80.8%; Pred. No. 5.3e-21;

Matches 173; Conservative 0; Mismatches 35; Indels 6; Gaps 1;

QY 1 TCAAGAGAGAGCGCTACGATCTGCGCGGCTGGGTGCTACAAAGTGAACAAGAGCTGG 60  
DB 1 TCAAGAGAGAGCGCTACGATCTGCGCGGCTGGGTGCTACAAAGTGAACAAGAGCTGG 60  
QY 61 GCGTGGCGCGCACCAATCCGCGCTCAGGTGACGACCAACCTCTCAGGAGGAAGCTGG 120  
DB 61 GCGTGGCGCGCACCAATCCGCGCTCAGGTGACGACCAACCTCTCAGGAGGAAGCTGG 114  
QY 121 TCGCCACCATCGATGCTGCTGCGCTGACAGAGGCGCCACGACGATGACCGCCCG 180  
DB 115 TCGCCACCATCGATGCTGCTGCGCTGACAGAGGCGCCACGATGAAGAGCTGG 174  
QY 181 GCGGCGTGCAGGTGCGGCTGATGTGACGACAT 214  
DB 175 GCGGCGTGCAGGTGCGGCTGATGTGACGACAT 208

RESULT 15

AAS05205 standard: DNA: 208 BP.

AAS05205;

07-SEP-2001 (first entry)

Mycoacterium tuberculosis rpoB gene fragment.

Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;  
PCR-restriction fragment length polymorphism analysis; ds.

Mycobacterium tuberculosis.

MO200131061-A1.

03-MAY-2001.

27-OCT-2000; 2000MO-KR01223.

PR 27-OCT-1999; 99KR-0046795.

PA (ERUM-) ERUME BIOTECH CO LTD.

PI Lee H, Park YK, Bal G, Kim S, Cho S, Kim Y, Park HJ;

DR WPI: 2001-300520/31.

XX New DNA fragments from the rpoB gene of mycobacteria, useful for  
PT diagnosis and identification of many mycobacterial species by  
restriction fragment length polymorphism

PS Disclosure: Page 41: 50pp: English.

XX The present sequence for Mycobacterium tuberculosis rpoB gene  
CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from  
CC various Mycobacterial species. These rpoB gene fragments can be used  
CC in the diagnosis and identification of Mycobacterium species using a  
CC novel PCR-restriction fragment length polymorphism analysis (PRA)  
CC method. The method comprises obtaining a restriction fragment length  
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,  
CC amplifying and digesting the DNA fragment from the microorganism to  
CC be identified and comparing the RFLP patterns from the known rpoB gene  
CC fragments with the unidentified fragment. The rpoB gene fragments  
CC are useful to identify a wide range of Mycobacterium species, e.g. for  
CC diagnosis or to obtain epidemiological and pathogenesis information for  
CC selection of appropriate therapies, including M. tuberculosis, M. leprae  
CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected  
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene  
CC fragments is rapid, precise, simple and cost effective (only 1 PCR  
CC experiment), and can differentiate between many species in a single  
CC test. Also described are oligonucleotide probes (AAS05227-AAS05242) for  
CC detecting specific Mycobacterial species.

SQ Sequence 208 BP: 46 A; 62 C; 67 G; 33 T; 0 other;

Query Match 60.6%; Score 129.6; DB 22; Length 208;  
Best Local Similarity 79.0%; Pred. No. 1.3e-19;

Matches 169; Conservative 0; Mismatches 39; Indels 6; Gaps 1;

QY 1 TCAAGAGAGAGCGCTACGATCTGCGCGGCTGGGTGCTACAAAGTGAACAAGAGCTGG 60  
DB 1 TCAAGAGAGAGCGCTACGATCTGCGCGGCTGGGTGCTACAAAGTGAACAAGAGCTGG 60  
QY 61 GCGTGGCGCGCACCAATCCGCGCTCAGGTGACGACCAACCTCTCAGGAGGAAGCTGG 120  
DB 61 GCGTGGCGCGCACCAATCCGCGCTCAGGTGACGACCAACCTCTCAGGAGGAAGCTGG 114  
QY 121 TCGCCACCATCGATGCTGCTGCGCTGACAGAGGCGCCACGACGATGACCGCCCG 180  
DB 115 TCGCCACCATCGATGCTGCTGCGCTGACAGAGGCGCCACGATGAAGAGCTGG 174  
QY 181 GCGGCGTGCAGGTGCGGCTGATGTGACGACAT 214  
DB 175 GCGGCGTGCAGGTGCGGCTGATGTGACGACAT 208

Search completed: November 12, 2002, 16:50:25  
Job time : 142.151 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 12, 2002, 16:50:46 ; Search time 654.37 Seconds

(without alignments)  
9517.553 Million cell updates/sec

Title: US-09-697-123b-21

Perfect score: 214

Sequence: 1 tcaagagaagcgtacgacat.....ccggtgcatgtgcagacat 214

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl.\*

1: gb\_ba.\*

2: gb\_hlg.\*

3: gb\_in.\*

4: gb\_om.\*

5: gb\_ov.\*

6: gb\_pat.\*

7: gb\_ph.\*

8: gb\_pl.\*

9: gb\_pt.\*

10: gb\_ro.\*

11: gb\_sts.\*

12: gb\_sy.\*

13: gb\_un.\*

14: gb\_vl.\*

15: em\_ba.\*

16: em\_fun.\*

17: em\_hum.\*

18: em\_in.\*

19: em\_mu.\*

20: em\_om.\*

21: em\_or.\*

22: em\_ov.\*

23: em\_pat.\*

24: em\_ph.\*

25: em\_pl.\*

26: em\_ro.\*

27: em\_sts.\*

28: em\_un.\*

29: em\_vl.\*

30: em\_hlg\_hum.\*

31: em\_hlg\_inv.\*

32: em\_hlg\_other.\*

33: em\_hlg\_mus.\*

34: em\_hlg\_pln.\*

35: em\_hlg\_rtd.\*

36: em\_hlg\_mam.\*

37: em\_hlg\_vrl.\*

38: em\_sy.\*

39: em\_hlg\_hum.\*

40: em\_hlg\_mus.\*

41: em\_hlg\_others.\*

Prod. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	132.8	62.1	3752	1	MSU24494
2	129.6	60.6	610	1	MTU318818
3	129.6	60.6	610	1	MTU318819
4	129.6	60.6	616	1	MTU318812
5	129.6	60.6	618	1	MTU318815
6	129.6	60.6	618	1	MTU318817
7	129.6	60.6	633	1	MTU318814
8	129.6	60.6	637	1	MTU318816
9	129.6	60.6	639	1	MTU318821
10	129.6	60.6	970	6	AX111339
11	129.6	60.6	3534	6	AX111339
12	129.6	60.6	3853	1	MTU12205
13	129.6	60.6	5084	1	MSGP08
14	129.6	60.6	19352	1	AE006964
15	129.6	60.6	19770	1	MTC1376
16	108.8	50.8	3447	6	AR067447
17	108.8	50.8	37617	1	MTB1790C
18	108.8	50.8	34850	1	MTB1790C
19	95.2	44.5	3941	1	MTB1790C
20	80.4	37.6	3495	6	AX120631
21	80.4	37.6	328050	1	AP005275
22	80.4	37.6	349980	6	AX127144
23	76	35.5	32923	1	SCD82
24	46.2	21.6	1672	1	RCPTPR
25	46	21.5	166050	1	AL646085
26	44.6	20.8	34023	1	SC2G2
27	44.4	20.7	190050	1	AL646059
28	44.2	20.7	8611	8	SCU74495
29	44	20.6	776	5	CHKTNC
30	44	20.6	39531	1	SCP8
31	43.8	20.5	1782	14	AF171937
32	43.8	20.5	11103	1	AE001944
33	43.2	20.2	1813	8	MZENRE
34	43	20.1	101365	2	AP004748
35	43	20.1	152263	2	AP005412
36	42.8	20.0	29625	1	SCN22A
37	42.8	20.0	77534	1	AF235504
38	42.6	19.9	24700	1	SCB8
39	42.6	19.9	113193	1	AF357202
40	42.6	19.9	133066	2	CNS08C80
41	42.4	19.8	10660	1	AE001963
42	42.2	19.7	1665	14	AF403050
43	42.2	19.7	1674	14	AF403049
44	42.2	19.7	1734	6	104901
45	42.2	19.7	1734	6	108509

#### ALIGNMENTS

RESULT 1

MSU24494

LOCUS MSU24494 3752 bp DNA linear BCT 02-MAR-2000

DEFINITION Mycobacterium smegmatis DNA polymerase (pob) gene, complete cds.

ACCESSION U24494

VERSION U24494.1 GI:790347

KEYWORDS

SOURCE

ORGANISM

Mycobacterium smegmatis.

Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;

Actinomycetales; Corynebacterineae; Mycobacteriaceae;

Mycobacterium.

1 (bases 1 to 3752)

REFERENCE

Hetherington, S.V., Watson, A.S. and Patrick, C.C.

Sequence and analysis of the pob gene of Mycobacterium smegmatis



Corynebacteriaceae; Mycobacteriaceae; Mycobacterium; Mycobacterium tuberculosis complex.

REFERENCE  
1  
AUTHORS  
TITLE  
Herrera, L., Jimenez, M.S. and Saez, J.A.  
Molecular analysis of rifampin-resistant Mycobacterium tuberculosis isolated in Spain (1996-2001). Description of new alleles into rpoB gene and review

JOURNAL  
REFERENCE  
2 (bases 1 to 610)  
AUTHORS  
TITLE  
Herrera, L.  
Direct Submission  
Submitted (07-AUG-2001) Herrera L., Servicio Bacteriología, Centro Nacional Microbiología, Ctra. Majadahonda-Pozuelo, Km 2.5, Majadahonda, Madrid. 28220, SPAIN

FEATURES  
source  
1. 610  
/organism="Mycobacterium tuberculosis"  
/isolate="1417-97"  
/db\_xref="taxon:1773"  
1. 610  
/gene="rpoB"  
/codon\_start=1  
/transl\_table=1  
/product="RNA polymerase beta subunit"  
/protein\_id="CAC87036.1"  
/db\_xref="GI:22208415"  
/translation="LDYRKLRGPEPTKESAOITLLENLFKEKRYDLARVGRYKVK  
KLGIVGEPTITSTLEEDVAVATLEYLVRLHGOITMTVPGVEVEVDIDHFGNR  
RLRTVGEIIONIRVMSRMERVRERMTQDVEATPOTLINIRPVVAIKKEFGTS  
QPSQFMQNNPLSGLTGHRRLSALGPGLSRERAGLEVDRVHP"

BASE COUNT  
122 a 191 c 202 g 95 t

ORIGIN

Query Match  
Best Local Similarity 60.68; Score 129.6; DB 1; Length 610;  
Matches 169; Conservative 0; Mismatches 39; Indels 6; Gaps 1;

Db  
1 TCAGAGAGAGCGCTACGATCTGCGCCGCTGCGTACAGAGTGAACAGAGCTGG 60  
|||||  
80 TCAGAGAGAGCGCTACGATCTGCGCCGCTGCGTACAGAGTGAACAGAGCTGG 139  
|||||  
61 GCGTGGCGGACCAATCCGGCTCAGGTGACACACACCTTCACCGAGAGAGCTGG 120  
|||||  
140 GCGTGGCGGACCAATCCGGCTCAGGTGACACACACCTTCACCGAGAGAGCTGG 193  
|||||  
121 TCAGAGAGAGCGCTACGATCTGCGCCGCTGCGTACAGAGTGAACAGAGCTGG 180  
|||||  
194 TGGCCACCATGATATCTGTCGCTTGCACAGAGGTGACACACGATGACCGTTCCGG 253  
|||||  
0Y 181 CCGGCGTGCAGGTCCGGTGCATGTGAGCAGACAT 214  
|||||  
Db 254 GCGGCGTGCAGGTCCGGTGCATGTGAGCAGACAT 287  
|||||

RESULT 4  
MTU318813 616 bp DNA circular BCT 09-AUG-2002  
LOCUS  
DEFINITION  
Mycobacterium tuberculosis partial rpoB gene for RNA polymerase  
beta subunit, isolate 1763-97.  
ACCESSION  
AJ318813  
VERSION  
AJ318813.1 GI:22208402  
KEYWORDS  
RNA polymerase beta subunit; rpoB gene.  
SOURCE  
Mycobacterium tuberculosis  
ORGANISM  
Mycobacterium tuberculosis  
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium  
tuberculosis complex.

REFERENCE  
1  
AUTHORS  
TITLE  
Herrera, L., Jimenez, M.S. and Saez, J.A.  
Molecular analysis of rifampin-resistant Mycobacterium tuberculosis isolated in Spain (1996-2001). Description of new alleles into rpoB

gene and review

JOURNAL  
REFERENCE  
2 (bases 1 to 616)  
AUTHORS  
TITLE  
Herrera, L.  
Direct Submission  
Submitted (07-AUG-2001) Herrera L., Servicio Bacteriología, Centro Nacional Microbiología, Ctra. Majadahonda-Pozuelo, Km 2.5, Majadahonda, Madrid. 28220, SPAIN

FEATURES  
source  
1. 616  
/organism="Mycobacterium tuberculosis"  
/isolate="1763-97"  
/db\_xref="taxon:1773"  
1. 615  
/gene="rpoB"  
/codon\_start=1  
/transl\_table=1  
/product="RNA polymerase beta subunit"  
/protein\_id="CAC87030.1"  
/db\_xref="GI:22208403"  
/translation="LDYRKLRGPEPTKESAOITLLENLFKEKRYDLARVGRYKVK  
KLGIVGEPTITSTLEEDVAVATLEYLVRLHGOITMTVPGVEVEVDIDHFGNR  
RLRTVGEIIONIRVMSRMERVRERMTQDVEATPOTLINIRPVVAIKKEFGTS  
QPSQFMQNNPLSGLTGHRRLSALGPGLSRERAGLEVDRVHP"

BASE COUNT  
125 a 191 c 201 g 99 t

ORIGIN

Query Match  
Best Local Similarity 60.68; Score 129.6; DB 1; Length 616;  
Matches 169; Conservative 0; Mismatches 39; Indels 6; Gaps 1;

Db  
1 TCAGAGAGAGCGCTACGATCTGCGCCGCTGCGTACAGAGTGAACAGAGCTGG 60  
|||||  
80 TCAGAGAGAGCGCTACGATCTGCGCCGCTGCGTACAGAGTGAACAGAGCTGG 139  
|||||  
61 GCGTGGCGGACCAATCCGGCTCAGGTGACACACACCTTCACCGAGAGAGCTGG 120  
|||||  
140 GCGTGGCGGACCAATCCGGCTCAGGTGACACACACCTTCACCGAGAGAGCTGG 193  
|||||  
121 TCAGAGAGAGCGCTACGATCTGCGCCGCTGCGTACAGAGTGAACAGAGCTGG 180  
|||||  
194 TGGCCACCATGATATCTGTCGCTTGCACAGAGGTGACACACGATGACCGTTCCGG 253  
|||||  
0Y 181 CCGGCGTGCAGGTCCGGTGCATGTGAGCAGACAT 214  
|||||  
Db 254 GCGGCGTGCAGGTCCGGTGCATGTGAGCAGACAT 287  
|||||

RESULT 5  
MTU318815 618 bp DNA circular BCT 09-AUG-2002  
LOCUS  
DEFINITION  
Mycobacterium tuberculosis partial rpoB gene for RNA polymerase  
beta subunit, isolate 2540-97.  
ACCESSION  
AJ318815  
VERSION  
AJ318815.1 GI:22208406  
KEYWORDS  
RNA polymerase beta subunit; rpoB gene.  
SOURCE  
Mycobacterium tuberculosis  
ORGANISM  
Mycobacterium tuberculosis  
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium  
tuberculosis complex.

REFERENCE  
1  
AUTHORS  
TITLE  
Herrera, L., Jimenez, M.S. and Saez, J.A.  
Molecular analysis of rifampin-resistant Mycobacterium tuberculosis isolated in Spain (1996-2001). Description of new alleles into rpoB  
gene and review

JOURNAL  
REFERENCE  
2 (bases 1 to 618)  
AUTHORS  
TITLE  
Herrera, L.  
Direct Submission  
Submitted (07-AUG-2001) Herrera L., Servicio Bacteriología, Centro

Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5,  
Majadahonda. Madrid. 28220, SPAIN

FEATURES  
source

1.618  
Location/Qualifiers

/organism="Mycobacterium tuberculosis"

/isolate="2540197"

/db\_xref="taxon:1773"

1.618  
/gene="rpoB"

<1..>618  
/gene="rpoB"

/codon\_start=1  
/transl\_table=11

/product="RNA polymerase beta subunit"

/protein\_id="CAC87032.1"

/db\_xref="GI:22208407"

/translation="TDEALLDLYRKLRPGEPTYESAQTLENIFFEKRYDLARYG  
YVNNKILHVEPTSSLTLEEDVVAIETLVLHSGOTMTVPGEVEPVEDID  
HFGNRLRTVGLIONOIRVGMSEMRVREMTODVEATPTLTINIRPVAAIKE  
FPGTSOLSGFMDONNPISGLTHKRRLLALGPGLSRERAGLEVVDY"

BASE COUNT 124 a 190 c 207 g 97 t

ORIGIN

Query Match

Best Local Similarity 79.0%; Score 129.6; DB 1; Length 618;  
Matches 169; Conservative 0; Mismatches 39; Indels 6; Gaps 1;

OY 1 TCAAGAGAAGCGCTACGATCTGGCCCGGCTGCTACAGAGTGAACAAGACCTCG 60

Db 95 TCAAGAGAAGCGCTACGATCTGGCCCGGCTGCTACAGAGTGAACAAGACCTCG 154

OY 61 GCCTGGCGCGCACCAATCCGGCTCAGGTGACCAACCAACCTCAGGAGAGAGCTCG 120

Db 155 GCCTGGCGCGCACCAATCCGGCTCAGGTGACCAACCAACCTCAGGAGAGAGCTCG 208

OY 121 TCGCCACATCGATGACCTGCTGGCCCGGCTGCTACAGAGTGAACAAGACCTCG 180

Db 209 TCGCCACATCGATGACCTGCTGGCCCGGCTGCTACAGAGTGAACAAGACCTCG 268

OY 181 GCGGCGTCCAGGTGCGCGGTGATGTGACGACAT 214

Db 269 GCGGCGTCCAGGTGCGCGGTGATGTGACGACAT 302

RESULT 6  
MTU318817 618 bp DNA circular BCT 09-AUG-2002

LOCUS Mycobacterium tuberculosis partial rpoB gene for RNA polymerase

DEFINITION beta subunit, isolate 2348-98.

ACCESSION AJ318817

VERSION AJ318817.1 GI:22208410

KEYWORDS RNA polymerase beta subunit; rpoB gene.

SOURCE Mycobacterium tuberculosis

ORGANISM Mycobacterium tuberculosis

REFERENCE 1

Herrera, L., Jimenez, M. S. and Saez, J. A.

Molecular analysis of rifampin-resistant Mycobacterium tuberculosis

isolated in Spain (1996-2001). Description of new alleles into rpoB

gene and review

Unpublished

2 (bases 1 to 618)

Herrera, L.

Direct Submission

Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro

Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5,

Majadahonda. Madrid. 28220, SPAIN

Location/Qualifiers

1.618  
/organism="Mycobacterium tuberculosis"

/isolate="2348-98"

/db\_xref="taxon:1773"

1.618  
/gene="rpoB"

<1..>618  
/gene="rpoB"

/codon\_start=1  
/transl\_table=11

/product="RNA polymerase beta subunit"

/protein\_id="CAC87034.1"

/db\_xref="GI:22208411"

/translation="TDEALLDLYRKLRPGEPTYESAQTLENIFFEKRYDLARYG  
YVNNKILHVEPTSSLTLEEDVVAIETLVLHSGOTMTVPGEVEPVEDID  
HFGNRLRTVGLIONOIRVGMSEMRVREMTODVEATPTLTINIRPVAAIKE  
FPGTSOPSPFGONNPISGLTHKRRLLALGPGLSRERAGLEVVDY"

BASE COUNT 124 a 192 c 207 g 95 t

ORIGIN

Query Match

Best Local Similarity 79.0%; Score 129.6; DB 1; Length 618;  
Matches 169; Conservative 0; Mismatches 39; Indels 6; Gaps 1;

OY 1 TCAAGAGAAGCGCTACGATCTGGCCCGGCTGCTACAGAGTGAACAAGACCTCG 60

Db 95 TCAAGAGAAGCGCTACGATCTGGCCCGGCTGCTACAGAGTGAACAAGACCTCG 154

OY 61 GCCTGGCGCGCACCAATCCGGCTCAGGTGACCAACCAACCTCAGGAGAGAGCTCG 120

Db 155 GCCTGGCGCGCACCAATCCGGCTCAGGTGACCAACCAACCTCAGGAGAGAGCTCG 208

OY 121 TCGCCACATCGATGACCTGCTGGCCCGGCTGCTACAGAGTGAACAAGACCTCG 180

Db 209 TCGCCACATCGATGACCTGCTGGCCCGGCTGCTACAGAGTGAACAAGACCTCG 268

OY 181 GCGGCGTCCAGGTGCGCGGTGATGTGACGACAT 214

Db 269 GCGGCGTCCAGGTGCGCGGTGATGTGACGACAT 302

RESULT 7  
MTU318814 633 bp DNA circular BCT 09-AUG-2002

LOCUS Mycobacterium tuberculosis partial rpoB gene for RNA polymerase

DEFINITION beta subunit, isolate 1058-97.

ACCESSION AJ318814

VERSION AJ318814.1 GI:22208404

KEYWORDS RNA polymerase beta subunit; rpoB gene.

SOURCE Mycobacterium tuberculosis

ORGANISM Mycobacterium tuberculosis

REFERENCE 1

Herrera, L., Jimenez, M. S. and Saez, J. A.

Molecular analysis of rifampin-resistant Mycobacterium tuberculosis

isolated in Spain (1996-2001). Description of new alleles into rpoB

gene and review

Unpublished

2 (bases 1 to 633)

Herrera, L.

Direct Submission

Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro

Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5,

Majadahonda. Madrid. 28220, SPAIN

Location/Qualifiers

1.633  
/organism="Mycobacterium tuberculosis"

/isolate="1058-97"

/db\_xref="taxon:1773"

1.633  
/gene="rpoB"

<1..>633  
/gene="rpoB"

/codon\_start=1

	Query Match	60.6%; Score 129.6; DB 1; Length 633;	
	Best Local Similarity	79.0%; Pred. No. 1.6e-12;	
	Matches 169:	Conservative 0; Mismatches 39; Indels 6; Gaps 1;	
QY	1 TCAAGGAGAAGCGCTACGATCTGGCCGCCCGTGGGTGCTACAGGTGAACAAGAACTGG 60		/transl_table=11 /product="RNA polymerase beta subunit" /protein_id="CAC87031.1" /db_xref="GI:22208405" /translation="DNTVGTDEALLDIYRKLRPESPPTKESAOTLLENLFPEKEKYDLARGRYKNKKILGLHVGEPITSTLTEDDVATLEYLRHLEGOTTMTVPGVGPVEATDDIDHEFNRLRTYGELIGNIRVGSMBRMYREBMTODVEATIPDTLINRPVVAAIKFEFTSOLSPFMDNNLSLTITYKRRLSALCPGLSNERAGLEVRY"
Dd	110 TCAAGGAGAAGCGCTACGATCTGGCCGCCCGTGGGTGCTATTAAGTTAACAAGAAGCTGC 169		
QY	61 GCCTTGCGCGGCACCAATCCGGCTCAGTGCACCACCAACCCCTCACCAGGAAGACGTGC 120		
Dd	170 GCCTCAGTCATGGCGGAGCC-----CATCACGCTCTCTCAGCCTGACCCGAAAGACGTGC 223		
QY	121 TCGCCACCATGAGTACTGCTGGCGCTGCACAGAGCGCCAGCACGATGACCCGCCCG 180		
Dd	224 TGCGCACCATTGAAATATCTGTGCCCTTGCAAGAGGGTCAAGACACGATGACCGTTCCGG 283		
QY	181 GCGGCGTCGAGTGCCTGGTGGATGGAGCGACAT 214		
Dd	284 GCGCGTCGAGTGCCTGGTGGAAACCGACGACAT 317		
RESULT 8	MTUJ18816	637 bp	DNA
LOCUS	MTUJ18816		circular BCF 09-AUG-2002
DEFINITION	Mycobacterium tuberculosis partial rpoB gene for RNA polymerase beta subunit, isolate 1255-96.		
ACCESSION	AJ318816		
VERSION	AJ318816.1 GI:22208408		
KEYWORDS	RNA polymerase beta subunit; rpoB gene.		
SOURCE	Mycobacterium tuberculosis.		
ORGANISM	Mycobacterium tuberculosis		
REFERENCE	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacteriaceae; Mycobacteriaceae; Mycobacterium; Mycobacterium tuberculosis complex.		
REFERENCE	1 Herrera, L., Jimenez, M. S. and Saez, J. A.		
AUTHORS	Molecular analysis of rifampin-resistant Mycobacterium tuberculosis isolated in Spain (1996-2001). Description of new alleles into rpoB gene and review unpublished		
TITLE	2 (bases 1 to 637)		
JOURNAL	Herrera, L.		
REFERENCE	Direct Submission		
AUTHORS	Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro Nacional Microbiologia, Ctra. Majadahonda-Pozuelo, Km 2.5,		
TITLE	Majadahonda, Madrid 28220, SPAIN		
JOURNAL	Location/Qualifiers		
FEATURES	1..637		
source	/organism="Mycobacterium tuberculosis"		
	/isolate="1255-96"		
	/db_xref="taxon:1773"		
gene	1..637		
CDS	/gene="rpoB"		
	<1..>637		
	/gene="rpoB"		
	/codon_start=1		
	/transl_table=11		
	/product="RNA polymerase beta subunit"		
	/protein_id="CAC87033.1"		
	/db_xref="GI:22208409"		
	/translation="VQTDALLADIYRKLRPESPPTKESAOITLLENLFPEKEKYDLARYGRIYKNKKILGLHVGEPITSTLTEDDVATLEYLRHLEGOTTMTVPGVGPVEATDDIDHEFNRLRTYGELIGNIRVGSMBRMYREBMTODVEATIPDTLINRPVVAAIKFEFTSOLSPFMDNNLSLTITYKRRLSALCPGLSNERAGLEVRY"		

BASE COUNT	128 a	198 c	210 g	101 t	
ORIGIN	IDHEFNRLRATVGLICNOIIVGMSRMEVRVREMTTODVEAIRPORTLINIRPVAA KEEFTSOLSOFMOPNNPLSGITKHKRMRFAIGPGLSERAGLEVRDHPSH*				
Query Match	60.6%	Score 129.6	DB 1	Length 637	
Best Local Similarity	79.0%	Pred. No. 1.6e-12			
Matches 169	Conservative	0	Mismatches 39	Indels	6
				Caps	1
Qy	1	TCGAAGAGAAAGGGCTACGATCTGGCCCGCGGTGGCTGCTACAGAGTGAACAAGACCTGG	60		
Db	101	TCGAAGAGAAAGGGCTACGACCTGGCCCGCGGTGGCTGCTATTAAGTCAACAAGAAAGCTTCG	160		
Qy	61	GGCTGGCGCGGACCAATCCGGCTCAGGTGACCCACCACCACTTCACGAGAAGACGTCG	120		
Db	161	GGCTGATGATCGGGGAGCC-----CATCAGCTCTGCACGCTGACCGAAGAACGCTCG	214		
Qy	121	TGGCCACCATTCAGTACTGCTGCCCTTCGACAGAGGGCCAGACCATGATGACCGCCCGC	180		
Db	215	TGGCCACCATTCGAAATATCTGGTCCCTTCGACAGAGGTGACACCATGATGACCGCTCCGG	274		
Qy	181	GGCGGCTGAGGTGCGCGGTGATGTGAGCGACAT	214		
Db	275	GGCGGCTGAGGTGCGCGGTGATGTGAGCGACAT	308		
RESULT 9					
MRUJ18821					
LOCUS	MRUJ18821	639 bp	DNA	circular	BCT 09-AUG-2002
DEFINITION	Mycobacterium tuberculosis partial rpoB gene for RNA polymerase				
ACCESSION	AJ318821	beta subunit, isolate 1071-98.			
VERSION	AJ318821.1	GI:22208418			
KEYWORDS	RNA polymerase beta subunit; rpoB gene.				
SOURCE	Mycobacterium tuberculosis.				
ORGANISM	Mycobacterium tuberculosis				
	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;				
	Corynebacteriineae; Mycobacteriaceae; Mycobacterium; Mycobacterium				
	tuberculosis complex.				
REFERENCE	1				
AUTHORS	Herrera, L., Jimenez, M. S. and Saez, J. A.				
TITLE	Molecular analysis of rifampin-resistant Mycobacterium tuberculosis				
	gene and review				
	isolated in Spain (1996-2001). Description of new alleles into rpoB				
	gene and review				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 639)				
AUTHORS	Herrera, L.				
TITLE	Direct Submission				
JOURNAL	Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro				
	National Microbiologia, Ctra. Majadahonda-Pozuelo, km 2.5,				
	Majadahonda. Madrid. 28220, SPAIN				
FEATURES	Location/Qualifiers				
source	1. 639				
	/organism="Mycobacterium tuberculosis"				
	/isolate="1071-98"				
	/db_xref="taxon:11773"				
gene	1. 639				
	/gene="rpoB"				
	<1..>639				
CDS	/gene="rpoB"				
	/codon_start=1				
	/transl_table=11				
	/product="RNA polymerase beta subunit"				
	/protein_id="CAC87038.1"				
	/db_xref="GI:22208419"				
	/translation="RRTDEALDIDYRKLPRGEPTRESAQTLLNLFERKRYDLARY				
	GRKVKVKKILHVGEPITSTLTEDVYATIEIVLHSGOTTMTVPQGVPEVETD				
	IDHFGRRRLRTGELIONIRVGMSEMEVVERBMTQVEAITTQTLINIPAAVAI				
	KEFGFSQLSQFMQNNPISGLTKQRRLALPGGLSRERAGLEVRDHPSHS"				
BASE COUNT	126 a	202 c	212 g	99 t	
ORIGIN					
Query Match	60.6%	Score 129.6	DB 1	Length 639	



Best Local Similarity 79.0%; Pred. No. 1,6e-12;  
Matches 169; Conservative 0; Mismatches 39; Indels 6; Gaps 1;

QY 1 TCAAGAGAAAGCGCTACGATCTGCCCCGCGTGGTGGTACACAGTGAACAGAAAGCTGG 60  
|||||  
Db 101 TCAAGAGAAAGCGCTACGATCTGCCCCGCGTGGTGGTACACAGTGAACAGAAAGCTGG 160  
61 GCTGGGCGGACCAATCCGGCTCAGGTGACCAACACACACCTCCACCGAGAGAGAGCTGG 120  
|||||  
Db 161 GCTGGGCGGACCAATCCGGCTCAGGTGACCAACACACACCTCCACCGAGAGAGAGCTGG 214  
QY 121 TCGCCACCATCAGTACCTGCTGCGCTGCACGAGGCGCACAGACAGATGACCGCCCG 180  
|||||  
Db 215 TGGCCACCATCAGTACCTGCTGCGCTGCACGAGGCGCACAGACAGATGACCGCTTCCG 274  
QY 181 GCGGCGTCGAGGTCGCGGTGATGTGACGACAT 214  
|||||  
Db 275 GCGGCGTCGAGGTCGCGGTGATGTGACGACAT 308

RESULT 10  
150706 150706 970 bp DNA linear PAT 07-OCT-1997  
LOCUS Sequence 1 from patent US 5643723.  
ACCESSION 150706  
VERSION 150706.1 GI:2472409  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 970)  
AUTHORS Persing, D.H., Hunt, J.J., Young, K.K.Y., Felmlie, T.A., Roberts, G.D.,  
and Whelan, A. Christian.  
TITLE Detection of a genetic locus encoding resistance to rifampin in  
mycobacterial cultures and in clinical specimens  
JOURNAL Patent: US 5643723-A 1 01-JUL-1997;  
FEATURES  
source 1..970  
location/Qualifiers  
BASE COUNT 182 a 302 c 330 g 156 t  
ORIGIN

Query Match 60.6%; Score 129.6; DB 6; Length 970;  
Best Local Similarity 79.0%; Pred. No. 1,5e-12;  
Matches 169; Conservative 0; Mismatches 39; Indels 6; Gaps 1;

QY 1 TCAAGAGAAAGCGCTACGATCTGCGCGGTGGTGGTACACAGTGAACAGAAAGCTGG 60  
|||||  
Db 26 TCAAGAGAAAGCGCTACGATCTGCGCGGTGGTGGTACACAGTGAACAGAAAGCTGG 85  
QY 61 GCTGGGCGGACCAATCCGGCTCAGGTGACCAACACACCTCCACCGAGAGAGAGCTGG 120  
|||||  
Db 86 GCTGGGCGGACCAATCCGGCTCAGGTGACCAACACACCTCCACCGAGAGAGAGAGCTGG 139  
QY 121 TCGCCACCATCAGTACCTGCTGCGCTGCACGAGGCGCACAGACAGATGACCGCCCG 180  
|||||  
Db 140 TGGCCACCATCAGTACCTGCTGCGCTGCACGAGGCGCACAGACAGATGACCGCTTCCG 199  
QY 181 GCGGCGTCGAGGTCGCGGTGATGTGACGACAT 214  
|||||  
Db 200 GCGGCGTCGAGGTCGCGGTGATGTGACGACAT 233

RESULT 11  
AX11339 3534 bp DNA linear PAT 30-APR-2001  
LOCUS Sequence 2072 from Patent W00123604.  
ACCESSION AX11339  
VERSION AX11339.1 GI:13927631  
KEYWORDS  
SOURCE Mycobacterium tuberculosis.  
ORGANISM Mycobacterium tuberculosis  
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;

Actinomycetales; Corynebacterineae; Mycobacteriaceae;  
Mycobacterium; Mycobacterium tuberculosis complex.  
REFERENCE 1 (bases 1 to 3534)  
AUTHORS Bergeron, M.G., Boissinot, M., Huletsky, A., m Nard, C., Ouellette, M.,  
Picard, F.J., and Roy, P.H.  
TITLE Highly conserved genes and their use to generate probes and primers  
for detection of microorganisms  
JOURNAL Patent: WO 0123604-A 2072 05-APR-2001;  
FEATURES  
source Infectio Diagnostic (I.D.I.) INC. (CA)  
location/Qualifiers  
1..3534  
/organism="Mycobacterium tuberculosis"  
/strain="Rv"  
/db\_xref="taxon:1773"  
BASE COUNT 679 a 1081 c 1188 g 586 t  
ORIGIN

Query Match 60.6%; Score 129.6; DB 6; Length 3534;  
Best Local Similarity 79.0%; Pred. No. 1,3e-12;  
Matches 169; Conservative 0; Mismatches 39; Indels 6; Gaps 1;

QY 1 TCAAGAGAAAGCGCTACGATCTGCGCGGTGGTGGTACACAGTGAACAGAAAGCTGG 60  
|||||  
Db 902 TCAAGAGAAAGCGCTACGATCTGCGCGGTGGTGGTACACAGTGAACAGAAAGCTGG 961  
QY 61 GCTGGGCGGACCAATCCGGCTCAGGTGACCAACACACCTCCACCGAGAGAGAGCTGG 120  
|||||  
Db 962 GCTGGGCGGACCAATCCGGCTCAGGTGACCAACACACCTCCACCGAGAGAGAGCTGG 1015  
QY 121 TCGCCACCATCAGTACCTGCTGCGCTGCACGAGGCGCACAGACAGATGACCGCCCG 180  
|||||  
Db 1016 TGGCCACCATCAGTACCTGCTGCGCTGCACGAGGCGCACAGACAGATGACCGCTTCCG 1075  
QY 181 GCGGCGTCGAGGTCGCGGTGATGTGACGACAT 214  
|||||  
Db 1076 GCGGCGTCGAGGTCGCGGTGATGTGACGACAT 1109

RESULT 12  
MTU12205 3853 bp DNA linear BCF 02-MAR-2000  
LOCUS Sequence 2072 from Patent W00123604.  
ACCESSION MTU12205  
VERSION MTU12205.1 GI:515684  
KEYWORDS  
SOURCE Mycobacterium tuberculosis.  
ORGANISM Mycobacterium tuberculosis  
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
Actinomycetales; Corynebacterineae; Mycobacteriaceae;  
Mycobacterium; Mycobacterium tuberculosis complex.  
REFERENCE 1 (bases 1 to 3853)  
AUTHORS Imboden, P., Troller, R., Marchesi, F., Telenti, A., Bodmer, T.,  
Cole, S., Schopfer, K., and Burkart, T.  
TITLE The rpoB gene of Mycobacterium tuberculosis  
JOURNAL Unpublished  
AUTHORS Imboden, P.  
REFERENCE 2 (bases 1 to 3853)  
TITLE Direct Submission  
JOURNAL Submitted (11-JUL-1994) Paul Imboden, Institute for Medical  
Microbiology, University of Berne, Friedbuehlstrasse 51, Berne,  
3010, Switzerland  
location/Qualifiers  
1..3853  
/organism="Mycobacterium tuberculosis"  
/strain="H37Rv"  
/db\_xref="taxon:1773"  
576..>3853  
/gene="rpoB"  
576..>3853  
/gene="rpoB"  
/codon\_start=1  
/transl\_table=11



KEYWORDS	Mycobacterium tuberculosis CDC1551.
SOURCE ORGANISM	Mycobacterium tuberculosis CDC1551. Bacteria: Firmicutes; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacteriinae; Mycobacteriaceae; Mycobacterium: Mycobacterium tuberculosis complex.
REFERENCE	1 (bases 1 to 19352)
AUTHORS	Fleischmann, R.D., Alland, D., Eisen, J.A., Carpenter, L., White, O., Peterson, J.J., DeBoy, R., Dodson, R., Gwinn, M., Haft, D., Hickey, E., Kolonyak, J.F., Nelson, W.C., Umayam, L.A., Ermolaeva, M., Salzberg, S.L., Delcher, A., Utterback, T., Weidman, J., Khouri, H., Gill, J., Mikula, A. and Bishai, W.
TITLE	Whole genome comparison of Mycobacterium tuberculosis clinical and laboratory strains
JOURNAL REFERENCE	Unpublished
AUTHORS	2 (bases 1 to 19352) Fleischmann, R.D., Alland, D., Eisen, J.A., Carpenter, L., White, O., Peterson, J.J., DeBoy, R., Dodson, R., Gwinn, M., Haft, D., Hickey, E., Kolonyak, J.F., Nelson, W.C., Umayam, L.A., Ermolaeva, M., Salzberg, S.L., Delcher, A., Utterback, T., Weidman, J., Khouri, H., Gill, J., Mikula, A. and Bishai, W.
TITLE	Direct Submission
JOURNAL	Submitted (25-APR-2001) The Institute for Genomic Research, 9712 Medical Center Dr. Rockville, MD 20850, USA
FEATURES	Location/Qualifiers
SOURCE	1. .19352

gene  
163..3699  
/gene-"MT0695"  
163..3699  
CDS  
/gene-"MT0695"  
/note-"similar to GB:L27989 GB:L05910 GB:U12205 SP:P47766  
PID:149992; Identified by sequence similarity; putative"  
/codon\_start=1  
/transl\_table=11  
/product-"DNA-directed RNA polymerase, beta subunit"  
/protein\_id="AAK44921.1"  
/db\_xref="GI:13880218"  
/translation="MLEGCIILDSROSKTASPSRPOSSNNNSYBAGNRYSPKAK  
REPLVYGLIYQNTSPFENLIGSPKMRSSAERDYNVPGLEVLYELSP IEDFSG  
MLSRSDPRFDVYKAPVDECKDKMTTAPLVFAETINNTGTEIKSIVFMDCFPP  
TEGFTI IGRKRVVAVSOLVNSPGVYFDETTIDKSTDTLSVYKIPSGAMLEFDV  
RDTVGRIIDKRRROPTVTLALAGTSSQIYERGSFSEIMKSTLEKNQNTGIDETAL  
IYRKILGPGEPPTKESAOQTLLENLFPEKERYDLAVGVYKVKIKGLVGEPTISL  
EDVAVTIEYVLRLEGGOTMTVPGCVYVPEVETDIDFGNRLRTVGTGLIONOYR  
MSSEVVEREMRTTODVEATIPOTLINRPVVALKEFEGFSOLSGOMDNQNLGSL  
HMRSLASLPGGLSREBRAGLEVROVHPYBRMKPIETPGPNLIGLISLVAUAPR  
LIEYIPRYKRVYDGVSDPEIYVITADEDDRVVAVQANSPIADAGRPEPRVLYRRKKA  
EVEYFSSSEVDYMDVSPROMSVATAMLEPLEHDDARMLKAMROAPVAVLSGA  
LVTEGHELMKAAIDAGDVVAEESGVIEVSADYITVMHDKGTRTYPMKRFASNGSA  
CANQC IYDAGGRVEAGV IADGPTCDGEMALGLIYLA IMPEHGNYIDA ILSN  
LVEEDVLTLS IIEHEEIDARDTKLGAETITPDIINISDEVLADIDEGVYICAEVR  
GILVQKVPKGETELPPEERILRLAIFEKAREVRDLSLVYHESGKXIGVIFSR  
DDEELPAGVNEFLRVYVAOKRKISDGDCLARHONKCVYICKILPEEMPLADGTGV  
IILNTHGVRPMNIGOLIECTLLPHMGCAHSGMKVDAKAPPMARILPELLEADQPNKI  
SNPVDSGAQAEALQGLISCTLPNRRGDVADNDKAMLPGRGSEPRPVYATGVYK  
MLTHNHYDKIHAIRSTGXPYSMTIQQPLGKAKORSGCRFMECHAMQATYATLDEE  
LTKISDITGVKVTIEATVKGENTPEPEIPSPFVLLKEUQSLCLNVEYLSGAAI  
LNEGEDDELERAAANLGINLSRNEASVEDLA"  
3744..7694  
/gene-"MT0696"  
3744..7694  
CDS  
/gene-"MT0696"  
/note-"similar to SP:P37871; identified by sequence  
similarity; putative"  
/codon\_start=1  
/transl\_table=11  
/product-"DNA-directed RNA polymerase, beta-prime subunit"  
/protein\_id="AAK44922.1"  
/db\_xref="GI:13880219"

/translation="MLDVNFDEBLICGLATVAEDIRONSYGCVKPEP1NRTLKPED  
 GLECKJEGPREDWECYCGYKRAVFKC1IECERGVEYTRAKVREREMGH1ELAPVAT  
 LHYMRGYSRGYLIDLPAPKDEK1IIFYAAYTVSDEERKHNEHSTLE2LEMAVYER  
 AVEDRDELEAKRAKLEADOLEAEACAKADARKRVADGGERMRO1RDAQOELDR  
 LEDISTESTLEKAPROLIVDENLYRVEYETPGAMAGSIOKLEIEND1DAEAS  
 LBDY1RNCKGKQKLBALRKLVAFAFOQSGNSPMGNULDANPV1PEPLRPNVOLDGR  
 PATSULND1YRVR1NRNKRRLIDULGPE1VNNRKMADSEVDALPDGCRKRPATV  
 GGNRNP1LSLDLKGQGRFRQNLKRVDSY1VGGPOLLKQHC1DQGRKPYALTE  
 LEKPEVMRLVDLNAHON1KSAKRAVEADRDQOMVH1YEAER1LNAP1LTHRG1  
 IOAFPEPMLVEGKAID1QLHVCFAENADDEGQOMVH1PL5MAEAKR1IMLSNN1L  
 SASAPR1PEMLDVTGLTYITTEFVEPDEYQSPASDHDEYQSSPAEM1AAR  
 GLSVRAIK1KRL1DQBRPVELEBELFESHMOMGDMMATTE1GTRFMPE1LPIG  
 FKNKMHKKVQA1IINDLAERYPMVAQY2DLKADGAFYATISQYTVSADYLVPP  
 RKKELLDHYEEBNADVEKOFORCALNMBERNBALVE1MKEATSDVEGVALRHE1PDDNP  
 IT1YDSGATGNPOTTR1LAQKGLNMBRKEF1IPRVCSSFEHGLVLEF1PHTHAA  
 RGLADT1LARTADSGYLRLVDVSDY1VIEHBCOCTERG1VLAERAPKAT1RPP  
 YLETSAYVARTLTDADGAGN1YVERGODDLEPDEIDALLAG1OQVARSVLTAT1ST  
 GCAICZYCRSMATKGLVD1GEAVNG1VMAOS1GEGEOT1DMTGFHOGVGD1TGGLRV  
 VOLE1EAPNRPKAP1ADVTRGVR1LEBGEORPCT1IYIPDQGGEEVGD1SKRORLAV  
 EKHGEGSRV1SDGHNVEYGQOLMBGASNDPHEYL1RVQCPREVO1HLVRVEQYTRAOG  
 VSIHKKHIEV1YRQMLRRV1IIDSSTFEFLGSL1DRAEF1EAENKRRV1AABGEPAAR  
 PLVKO1ITASLTADSWLSAASFOETTR1VLTPAALINCRDK1AGNRRV1IGK1LPACT  
 G1NR1RNA1NOPTLEBARAAAY1TIPSYEDQYYSDFGCAATGAAY1LDDYGSYSR"  
 complement("7691", .8065)  
 /gene="MT0697"  
 complement("7691", .8065)  
 /gene="MT0697"

gene  
/product="hypothetical protein"  
/protein\_id="AAK44923.1"  
/db\_xref="GI:13880220"  
/translation="MFDSAAAITNGHAMASAMERSGLIECVAGLIDGPEGFADKL  
NEDRSRRVRPRQADGCIATHVERGGGROSGOAGVYPRMHGFALAMODLIHH  
GEOTONRIAOAFRVYFCVCSPT"  
complement(8058..9972)  
/gene="MT06698"  
/note="this region contains an authentic frame shift and  
is not the result of a sequencing artifact; Identified by  
glimmer2; putative; conserved hypothetical protein,  
authentic frameshift"  
10167..10925  
/gene="MT06699"  
10167..10925  
/gene="MT06699"  
/note="identified by match to PFAM protein family HMM  
PF01261"  
/codon\_start=1  
/transl\_table=1  
/product="AP endonuclease, family 2"  
/protein\_id="AAK44924.1"  
/db\_xref="GI:13880221"  
/translation="MLIGSHVPTDPLAAAEAGADVVOIFLGNOSMKAPKPPDMA  
ALKATILTYHAPYLINLIASANNRVRIPSRKTIQFCAAADTGAAYVHGCHVD  
DNDIDCGORRKALDRLETEVPLEYLTAGGDHMAARRPDTIRLMDVIGDGTGIC  
LDCTMAAGALTDADVDRITKAITGRIDLVHCNDSRODAGSGRDRHNLDSGOIDPDL  
LVAAVKAAGAVPCTADQGRKDIADFLERTGS"  
10957..11799  
/gene="MT0700"  
10957..11799  
/gene="MT0700"  
/note="similar to GB:U00012.PID:468683; Identified by  
sequence similarity; putative"  
/codon\_start=1  
/transl\_table=1  
/product="hydrolase/esterase, putative"  
/protein\_id="AAK44925.1"  
/db\_xref="GI:13880222"  
/translation="MLRRVAILLAALVLAFAAGCGGTRLAAGFNGNSVHTLDDVDAGR  
SYRLKPLGLSSAPLVMLHGGFGSAKOERSYGMDELADSEKFLVAYPDGYHAWN  
ANGGCCGRPRECVDDIGFRAVAVDIAINNVSDPAIRVTCMSGAINSTLPDNT  
SIPATGVSTQTLDPGOSPPRVITIRHCRADLVYHCGGRCGRARINCPDQNT

gene  
CDS  
11859..13487  
/gene="MT0701"  
/note="similar to SP:P33224 GB:L20915 PID:457172  
PID:457174 PID:537028; identified by sequence similarity;  
putative"

gene  
CDS  
13498..14436  
/gene="MT0702"  
13498..14436  
/gene="MT0702"  
/note="similar to GP:3885480; identified by sequence  
similarity: putative"  
/codon\_start=1  
/transl\_table=11  
/product="acyl-CoA dehydrogenase, putative"  
/protein\_id="AAK44926.1"  
/db\_xref="GI:13880223"  
/translation="MSDHYVTVNOVPLENYNPASSPVLTALIOEGCGMDDEVNY  
GAISASCOBMRGELADNRNPLTHHDAHYRVDVEYDPRYHELMKTAITHGHAAP  
MADRRGAVHVAKTSYVWYERGHICISMTYAVPALRINSELAAYKPEPLTSREY  
DDELPAATTKAGITGKSMTEKOGSSDVRAGTQATPADGSYSTLGHKFTSAPMD  
ITFLVAQAPDGLSCFLPRVLPDGTNRNMFQRLDKLGNHANSSEVEYDAVAMLY  
GEGRGVPTIIEVNLTRLDALGATSMTGLRAVHAQHRKAFGAYLIDOPLMRN  
VLADLAVEAEATIVAMRAGATONAVRGNETEALRRIGLAALKYVWCKRSTHAHE  
ALECLGNGYVEDSGMPRLYREAPLMGIMEGSGVSLDITLRAMTRPACVEVLFDEL  
ARAGODPRLDHYRLRPLQDGLDTGYRARKIAEDICIALGSLLVRRGHFAVMEA  
FLATRLGGOMGAGTMAPGLADLAPILERALVKG"

gene  
CDS  
14439..15161  
/gene="MT0703"  
14439..15161  
/gene="MT0703"  
/note="identified by Glimmer2; putative"  
/codon\_start=1  
/transl\_table=11  
/product="hypothetical protein"

Query Match 60.6% Score 129.6; DB 1; Length 19352;  
Best Local Similarity 79.0% Pred. NO. 1e-12;  
Matches 169; Conservative 0; Mismatches 39; Indels 6; Gaps 1;

QY 1 TCACGAGAACGCTACGATCTGCGCGCGTGGTACGACGACGACGACGCTG 60  
DB 1064 TCACGAGAACGCTACGATCTGCGCGCGTGGTACGACGACGACGACGCTG 1123  
QY 61 GCGTGGCGGACCAATCGGCTCAGTGACACACCTCACCAGGAGAGAGCTCG 120  
DB 1124 GCGTGGCGGACCAATCGGCTCAGTGACACACCTCACCAGGAGAGAGCTCG 1177  
QY 121 TCGCCACCATGAGTACCTGCTGCGCGCGGCGGACGACGACGACGACGACGCTG 180  
DB 1178 TCGCCACCATGAGTACCTGCTGCGCGCGGCGGACGACGACGACGACGACGCTG 1237  
QY 181 GCGGCGTGAAGTCCGCTGAGTGTGAGACACAT 214  
DB 1238 GCGGCGTGAAGTCCGCTGAGTGTGAGACACAT 1271

RESULT 15  
MTC1376

LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
MTC1376 19770 bp DNA linear BCT 03-AUG-2001  
Mycobacterium tuberculosis H37Rv complete genome; segment 32/162.  
295972.1 AL123456  
295972.1 GI:3261190

REFERENCE  
AUTHORS  
1 (bases 1 to 19770)  
Cole,S.T., Brosch,R., Parkhill,J., Garnier,T., Churcher,C.,  
Harris,D., Gordon,S.V., Eiglmeier,K., Gas,S., Barry III,C.E.,  
Tekala,F., Badcock,K., Basham,D., Brown,D., Chillingworth,T.,  
Connor,R., Davies,R., Devlin,K., Fellwell,T., Gentles,S.,  
Hamlin,N., Holroyd,S., Hornsby,T., Jagels,K., Krogh,A., McLean,J.,  
Moule,S., Murphy,L., Oliver,S., Osborne,J., Quail,M.A.,  
Rajandream,M.A., Rogers,J., Rutter,K., Seeger,K., Skelton,S.,  
Squires,S., Squires,R., Sulston,J.E., Taylor,K., Whitehead,S. and  
Barrell,B.G.

TITLE  
JOURNAL  
MEDLINE  
PUBMED  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
2 (bases 1 to 19770)  
Parkhill,J.  
Submitted (11-JUN-1998) Submitted on behalf of the Mycobacterium  
tuberculosis sequencing and mapping teams, Sanger Centre, Wellcome  
Trust Genome Campus, Hinxton, Cambridge CB10 1SA Unité de Génétique  
Moléculaire Bactérienne, Institut Pasteur, 28 rue du Docteur Roux,  
75724 Paris Cedex 15, France E-mail: parkhill@sanger.ac.uk  
On Jun 27, 1998 this sequence version replaced gi:2143285.

COMMENT  
Details of M. tuberculosis sequencing at the Sanger Centre are  
available on the World Wide Web.  
(URL: <http://www.sanger.ac.uk/projects/M.tuberculosis/>) CDS have  
been renumbered from the original cosmid submissions but the old  
gene designations are in brackets after the new gene numbers.  
Gene prediction was based on a Hidden Markov Model of TB genes.  
Implemented in TParse (Krogh) supplemented with visual inspection  
of positional base preference in codons, especially where there is  
an increase in the observed/expected third position G + C.  
CAUTION: In some cases we may not have predicted the correct  
initiation codon. Where possible we choose an initiation codon  
(atg, gtg, or ttg) which is preceded by an upstream ribosome  
binding site sequence (optimally 5-13bp before the initiation  
codon). If this cannot be identified we choose the most upstream  
initiation codon.

FEATURES  
SOURCE  
1..19770  
/organism="Mycobacterium tuberculosis H37Rv"  
/strain="H37Rv"  
/db\_xref="taxon:83332"  
/clone="1376"  
57..61  
/note="tagga, possible rbs upstream of Rv0654"  
68..1573  
/gene="Rv0654"  
68..1573  
/note="Rv0654"  
/note="Rv0654" (MTC1376.22), len: 501. unknown, FASTA  
score: 0.5353 LIGNOSTILBERE-ALPHA,BETA-DIOXGENASE (485  
aa) opt:280 z-score: 330.1 E(1): 2.3e-11, (28.5% identity  
in 523 aa overlap). Also similar to M. tuberculosis protein  
MTCY21C12.07c (29.5% identity in 522 aa overlap)."  
/codon\_start=1  
/transl\_table=11  
/product="hypothetical protein Rv0654"  
/protein\_id="CA009380.1"  
/db\_xref="GI:2143307"  
/db\_xref="SPTREMBL:006785"



Wed Nov 13 13:45:44 2002

us-09-697-123b-21.rge

Page 11

Db 10932 GCGGCTCGAGTGCCTGGGAACGACGACAT 10965

Search completed: November 13, 2002, 01:27:35  
Job time : 679.07 secs



/organism="Mycobacterium smegmatis"  
/strain="mc2155"  
/db\_xref="taxon:1772"  
/clone\_lib="Mycobacterium anaerobic stationary phase  
library"  
/lab\_host="E. coli XL1-Blue MRF"  
/note="Vector: Lambda ZAP II; BamCI were disrupted using  
a Mini bead beater (Biospec). RNA was isolated using the  
RNeasy protocol (Qiagen). Purified RNA was subjected to 2  
rounds of digestion with RNase-free DNase I (Promega).  
DNase I was heat-inactivated at 75°C for 5 min. and  
extraction and ethanol precipitation. The RNA  
preparations were confirmed to be free of genomic DNA  
contamination by carrying out PCR and RT-PCR using the  
histone-like protein gene hlp (Lee et al., 1998). cDNA  
was synthesized using random hexamer primers (Promega)  
and Stratagene's cDNA synthesis kit. cDNA fragments were  
ligated into Lambda ZAP II vector and packaged in vitro  
using Stratagene reagents."

BASE COUNT 230 a 395 c 405 g 252 t  
ORIGIN

Query Match 83.8%; Score 174.4; DB 9; Length 1282;  
Best Local Similarity 89.9%; Pred. No. 1,1e-28;  
Matches 187; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 1 TCACGAGAGAGCGCTACGACCTGGCCGCGCCGCTACAGGTACACAGAGCTGG 60  
DB 236 TCACGAGAGAGCGCTACGACCTGGCCGCGCTTACAGGTACACAGAGCTGG 177  
QY 61 GCCTGACGCGCGCCGATCAGCTGCTGCTGACGAGAGAGAGCTGCTGCCA 120  
DB 176 GCCTGACGCGCGCCGATCAGCTGCTGCTGACGAGAGAGAGCTGCTGCCA 117  
QY 121 CCATGAGTACTGCTGGCGCTGCACGAGCCAGACAGATGACCTCCCGCGCGC 180  
DB 116 CCATGAGTACTGCTGGCGCTGCACGAGCGGTGACACTTGATGACCTCCCGCGCGC 57  
QY 181 TCGAGTCCCGGTGAGGTGAGACAT 208  
DB 56 TCGAGTCCCGGTGAGGTGAGACAT 29

RESULT 2  
BI995346  
LOCUS 568 bp mRNA linear EST 25-OCT-2001  
DEFINITION 1031026E07.Y2 C. reinhardtii CC-1690, Stress II (normalised),  
Lambda zap II Chlamydomonas reinhardtii cDNA, mRNA sequence.  
ACCESSION BI995346  
VERSION BI995346.1 GI:16430136  
KEYWORDS EST.  
SOURCE Chlamydomonas reinhardtii.  
ORGANISM Chlamydomonas reinhardtii  
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;  
Chlamydomonadales; Chlamydomonas.  
REFERENCE 1 (bases 1 to 568)  
Grossman, A., Chan, C.-W., Davies, J., Harris, E., Hauser, C., Lefebvre  
P., McDermott, J.P., Shrager, J., Silflow, C. and Stern, D.  
Analyses of the Chlamydomonas reinhardtii Genome: A Model,  
Unicellular System for Analyzing Gene Function and Regulation in  
Vascular Plants. Project: 1031  
JOURNAL Unpublished (2001)  
COMMENT Contact: Charles Hauser  
DCMB Box 91000  
Duke University  
Durham, NC 27708-1000  
Tel: 919 613 8159  
Fax: 919 613 8177  
Email: chauser@duke.edu  
FEATURES  
source 1..568 Location/Qualifiers

/organism="Chlamydomonas reinhardtii"  
/strain="CC-1690 wild type mt+ 21gr"  
/db\_xref="taxon:3055"  
/clone\_lib="C. reinhardtii CC-1690, Stress II (normalised  
)", Lambda ZAP II"  
/note="Vector: pBluescript II SK-; Site 1: EcoRI; Site 2:  
XhoI; Stress condition II library, constructed by John  
Davies and Jeffrey McDermott, combines cDNAs from CC-1690  
cells grown to mid-log phase in TAP (NH4+ containing)  
and shifted to TAP -N03- (24hrs); H2 production  
conditions (0, 12hr, 24hr) see Wells et al., (2000) Plant  
Phys. 122: 127-135; TAP + H2O2 (1, 12, 24 hr); TAP +  
sorbitol (1, 2, 6, 24 hr); TAP + Cd (1, 2, 6, 24 hr).  
PolyA mRNA was purified from each sample, pooled and cDNA  
synthesized. The cDNA was directionally cloned into Lambda  
ZAP II (Stratagene) in the EcoRI (5') and XhoI (3')  
sites. pBluescript II SK- plasmids were excised from the  
lambda ZAP clones by superinfection with Exsist  
(Stratagene) phage. The library was normalized using  
method 4 described in Bonaldo et al., (1996) Genome  
Research 6: 791-806."

BASE COUNT 80 a 206 c 185 g 97 t  
ORIGIN

Query Match 24.1%; Score 50.2; DB 13; Length 568;  
Best Local Similarity 52.7%; Pred. No. 0.16;  
Matches 109; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

QY 2 CAAGAGAGAGCGCTACGACCTGGCCGCGCCGCTACAGGTACACAGAGCTGG 61  
DB 251 CAAGAGAGAGCGCTTCCAGAGTGTGACGCGCTTGAACAGGTGATGAGGTGG 310  
QY 62 CCGTACGCGCGCCGATCAGCTGCTGCTGACGAGAGAGAGCTGCTGCCA 121  
DB 311 CGTGGCTGACGAGGAGCAGCGCGCGCTCTCGCGCGCTGCTTACACTGGCGCC 370  
QY 122 CATGAGTACTGCTGGCGCTGCACGAGCGCCAGACAGATGACCTCCCGCGCGCT 181  
DB 371 CGCGCGCGACCGCCGCTGTTCCCGCGCGCGCGGTGTCGCTCCACTGCGCG 430  
QY 182 CGAGTCCCGGTGAGGTGAGACAT 208  
DB 431 CGACAGCGCGGTGAGGTGAGACAT 457

RESULT 3  
AO687771/c  
LOCUS 877 bp DNA linear GSS 01-JUL-1999  
DEFINITION nbxb0075L01f CUGI Rice BAC Library Oryza sativa genomic clone  
ACCESSION AO687771  
VERSION AO687771.1 GI:5328939  
KEYWORDS GSS.  
SOURCE Oryza sativa.  
ORGANISM Oryza sativa  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Eriaristidae; Oryzaceae; Oryza.  
REFERENCE 1 (bases 1 to 877)  
Wing, R.A. and Dean, R.A.  
A BAC End Sequencing Framework to Sequence the Rice Genome  
Unpublished (1998)  
JOURNAL Contact: Ming RA  
Clemson University Genomics Institute  
100 Jordan Hall, Clemson, SC 29634, USA  
Tel: 864 656 7288  
Fax: 864 656 4293  
Email: rwing@clemson.edu  
Seq primer: TAATCGACTCAGCTATAGCG  
Class: BAC ends  
FEATURES  
High quality sequence stop: 85.  
Location/Qualifiers





```

/tissue_type="Roots"
/dev_stage="Full tillering"
/lab_host="E. coli SOLR"
/note="Vector: Lambda uni-zap XR, excised phagemid
bluescript SK(-); Site_1: EcoRI; Site_2: XhoI; Hydropenic
plants grown to full tillering stage were treated with 150
mM NaCl for either 12 hours or 7 days. Root tissues of the
plants subjected to both types of treatment were collected
separately at University of California, Davis (E. Akhnov
and K. Deal in J. Dvorak's lab). Total RNA was prepared
separately from the two samples (12h and 7day treatments),
and equal amount of RNA was then pooled. PolyA RNA was
purified from the pooled RNA, a cDNA library was made, and
the cDNA clones were in vivo excised to give bluescript
SK(-) phagemids in J. Dvorak's lab (E. Akhnov, J. Dvorak)
at the University of California, Davis. Colony plating,
plasmid DNA preparations and DNA sequencing were performed
in the OD Anderson lab (all other authors)."
BASE COUNT      92 a      198 c      130 g      142 t
ORIGIN
Query Match      21.3%; Score 44.2; DB 14; Length 562;
Best Local Similarity 54.7%; Pred. No. 3.3;
Matches 88; Conservative 0; Mismatches 73; Indels 0; Gaps 0;
QY 25 CCGCGGTGGCGGCTACAAAGCTGCAAGAGCTGGCGCTGAACGCCGCCAGCCGATCA 84
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 160 CCGTCACTGCTCCGCCGCCGCTCAACATCTTCAAGAGCGCTCAACATCTCAAGCCGCCCA 219
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 85 CGTCGCTGACTCTGACCGAGAGACGCTGCGCCACCATGAGTACCTGCTGCGCTGC 144
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 220 CCGGCTGGAAGACGACGACGTCATCTCGCAGCCGCTCGCGGCTGCTCTCC 279
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 145 ACGAGGGCCAGACACGATGACCGTCCCGCGGCGCTGAG 185
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 280 TCGAGCGCATCATCGTGGCGCTGCTGCGCGCGCCGCAAG 320
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 6
LOCUS      BE227740      370 bp      mRNA      linear      EST 06-JUL-2000
DEFINITION      894034G12.y1 C. reinhardtii CC-1690, normalized, Lambda zap II
ACCESSION      BE227740
VERSION      BE227740.1
KEYWORDS      GI:8932979
SOURCE      EST.
ORGANISM      Chlamydomonas reinhardtii.
                Chlamydomonas reinhardtii
                Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
                Chlamydomonadaceae; Chlamydomonas.
REFERENCE      1 (bases 1 to 370)
AUTHORS      Grossman, A., Davies, J., Federspiel, N., Harris, E., Lefebvre, P.,
                McDermott, J.P., Silflow, C., Stern, D., and Surzycki, R.
TITLE      Analysis of the Chlamydomonas reinhardtii Genome: A Model,
                Unicellular System for Analyzing Gene Function and Regulation in
                Vascular Plants; project phase 2
JOURNAL      Unpublished (2000)
COMMENT      Contact: Elizabeth H. Harris
                DCMB Box 91000
                Duke University
                Durham, NC 27708-1000, USA
                Tel: 919 613 8164
                Fax: 919 613 8177
                Email: chlamy@duke.edu.
FEATURES
source      1..370
                /organism="Chlamydomonas reinhardtii"
                /strain="CC-1690 wild type mt+ 21gr"
                /db_xref="taxon:3055"
                /clone_lib="C. reinhardtii CC-1690, normalized, Lambda zap
                II"
                /note="Vector: Bluescript II SK(-); Site_1: EcoRI; Site_2:
                XhoI; This library, constructed by John Davies and Jeffrey

```

```

McDermott, combines cDNAs from CC-1690 cells grown to
mid-log phase in YAP (acetate-containing) medium in the
light, YAP medium in the dark, HS (minimal) medium in
ambient levels of CO2 and HS medium bubbled with 5% CO2.
PolyA mRNA was purified from each sample, pooled and cDNA
synthesized. The cDNA was directionally cloned into lambda
zap II (Stratagene) in the EcoRI (5') and XhoI (3') sites.
bluescript II SK- plasmids were excised from the lambda
zap clones by superinfection with Exsist (Stratagene)
phase. The library was normalized using method 4 described
in Bonaldo et al (1996) Genome Research 6: 791-806."
BASE COUNT      68 a      123 c      119 g      60 t
ORIGIN
Query Match      20.9%; Score 43.4; DB 10; Length 370;
Best Local Similarity 50.7%; Pred. No. 4.8;
Matches 104; Conservative 0; Mismatches 101; Indels 0; Gaps 0;
QY 2 CAAAGAGAGCGCTACAGCTGGCGCGCTGGCGCCCTCAAGGTCACAGAGCTGGG 61
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 113 CAGCCACAGACCCCGCGCGCGGAGAACGACTTCGGCATCAAGTTCACATCAACAGCGG 172
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 62 CTTGAAGCGCGCGCCAGCCCATCAGCTGCTGACTGTGACCGAGAGAGAGCTGTGCCAC 121
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 173 CGAGCCCGCGCGCGAGCGCATCACCAAGATCTTTGGCGAGACCTCCAGAGTGGCCAC 232
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 122 CATCGAGTACTGTGCTGCGCTGCAGGAGGCCAGACAGATGACCTCCCGCGGCGCT 181
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 233 GCTCAACATGCGCGACATCCCGAGCGTCGACTGTGGAAGTGGCGGTACCAAGTTCGG 292
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 182 CGAGGTCCCGGTGAGAGTGAGCAGC 206
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 293 CGAGTTTCAGAGTGGAGTGTGTGAC 317
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 7
LOCUS      B0767689      391 bp      mRNA      linear      EST 26-JUL-2002
DEFINITION      EBr008_SQ009_H13.R root, 3 week, drought-stressed, cv Optic, EBr008
ACCESSION      B0767689
VERSION      B0767689.1
KEYWORDS      GI:21976163
SOURCE      EST.
ORGANISM      Hordeum vulgare.
                Hordeum vulgare
                Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae
                ; Triticeae; Hordeum.
REFERENCE      1 (bases 1 to 391)
AUTHORS      Hedley, P., Liu, H., Caldwell, D., McCallum, N., Mudie, S., Cardle, L.,
                Ramsay, L., Machray, G., Marshall, D.F.M., and Maugh, R.
TITLE      Development of Barley Transcriptome Resources
JOURNAL      Unpublished (2001)
COMMENT      Contact: Maugh R. Marshall DF
                Genome Dynamics/Computational Biology
                Scottish Crop Research Institute
                Invergowrie, Dundee, DD2 5DA, Scotland, UK
                Tel: 00 44 1382 562731
                Fax: 00 44 1382 562426
                Email: estescri.sari.ac.uk.
FEATURES
source      1..391
                /organism="Hordeum vulgare"
                /cultivar="Optic"
                /db_xref="taxon:4513"
                /clone_lib="EBr008_SQ009_H13"
                /clone_lib="root, 3 week, drought-stressed, cv Optic,
                EBr008"
                /tissue_type="root"
                /dev_stage="3 week"
                /lab_host="DH10B"
                /note="Vector: pSPORT1; Site_1: Sal I; Site_2: Not I;
                Non-normalised library, directionally cloned into pSPORT1.

```





QY	1	TCAGGAGAGGCGTACAGCACTGGCCCGCGTGGCGGCGCTACAGAGTCAACAAGAGCTGG	60
	104	TCACGCGGAATACAGCACTCTCTGCTCCGACCCAGCGCTCCACATCGCTGGCAGAGTGAGC	163
QY	61	GCCTGAAGCGCGGACGCGCATCGACGTCGTCGACGCTGACGAGGAGAAAGCTGTCGCCA	120
Db	164	TCGCCGTGCACACACTGCTCTCTACGCTCCCGGCGCTGGCAAGAGAAATACACCCGCG	223
QY	121	CCATGAGTACCTGTCGTGGCGCTGCACAGAGGCGGACAGACACATGACGCTCCCGCGGCG	180
Db	224	TCATAGGACGACCGCAGGCGCTGGCGAGTGGCAGACACATCACCCGATGGGCTCTA	283
QY	181	TCGAGTCCCGGTGGAGTGGAGACA	207
Db	284	ACGTGCTGCGCGAGGCGCTTCGACACA	310
RESULT 13			
LOCUS	AY108449	1072 bp	mRNA
DEFINITION	Zea mays PCO131734 mRNA sequence.		
ACCESSION	AY108449		
VERSION	AY108449.1	GI:21211527	
KEYWORDS	HTC.		
SOURCE	Zea mays.		
ORGANISM	Zea mays.		
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophytes; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC		
	Clade; Panicoideae; Andropogoneae; Zea.		
REFERENCE	1 (bases 1 to 1072)		
AUTHORS	Hallney,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S., Athur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.		
TITLE	Maize Mapping Project/Dupont Consensus Sequences for Design of Overgo Probes		
JOURNAL	Unpublished (2002)		
REFERENCE	2 (bases 1 to 1072)		
AUTHORS	Coe,E.C.		
TITLE	Direct Submission		
JOURNAL	Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA		
FEATURES	Location/Qualifiers		
source	1..1072		
	/organism="Zea mays"		
	/db_xref="MaizeDB:637686"		
	/db_xref="taxon:4577"		
	/clone="PCO131734"		
	/clone_lib="Maize Mapping Project/Dupont Consensus Library"		
	/note="this sequence is part of a project of EST assemblies resulting from the application of public contigs to seed Dupont contigs; this resource was assembled by Dupont as part of a collaboration for the overgo addressing of BACs in conjunction with the Maize Mapping Project"		
BASE COUNT	216 a 348 c 318 g 190 t		
ORIGIN			
Query Match	20.3%:	Score 42.2:	DB 11: Length 1072:
Best Local Similarity	51.9%:	Pred. No. 9.6:	
Matches	95: Conservative	0: Mismatches	88: Indels 0: Gaps 0:
QY	21	CTGCGCCGCGTGCGCCCTACAGAGTCAACAGAGCTGGGCTGAAGCGCGGACGCG	80
Db	300	CTGATCAAGCTAGGCGGCGCGCGGCGGCGGTGAGTGGGCTCTTTCACGGGCTACTCG	359
QY	81	ATACGTCGTGCACGCTGACCGAGAGAGCTGTCGCCACCATCATGACCTGTCGGC	140
Db	360	CTGCTGGCAGCGGCTGTGCGCTCCCGCGAGCGCAAGGTATGCAATTCGACGTGAGC	419
QY	141	CTGACAGAGGCGCAGACCATGATGACGTCCCGCGGCGGTGAGTGTCCGCTGAGGTG	200
Db	420	CGGAGTACTACGACATGCGGCGCGCCCTTATCGACGCGCGCGGAGTGGCAAGGTG	479

OY 201 GAC 203  
DB 480 GAC 482

# RESULT 14

LOCUS BG315577 453 bp mRNA linear EST 28-NOV-2001

DEFINITION SabH109.y1 Gm-cl032 glycine max cDNA clone GENOME SYSTEMS CLONE

ACCESSION BG315577  
VERSION BG315577.1 GI:13125034

KEYWORDS soybean, glycine max

ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine

REFERENCE 1 (bases 1 to 453)

## AUTHORS

Shoemaker, R., Keim, P., Vodkin, L., Erpelting, J., Corvett, V., Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.

TITLE Public Soybean EST Project  
JOURNAL Unpublished (1999)  
COMMENT Contact: Shoemaker R/Public Soybean EST Project  
Public Soybean EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@wustl.edu

This clone is available through: Resgen, Invitrogen Corp. 2130 South Memorial Parkway Huntville, AL 35801 For further information call: (800)-533-4363 or contact via email: ccuteresgen.com  
High quality sequence stop: 180.

## FEATURES

Location/Qualifiers

1..453  
/organism="Glycine max"  
/db\_xref="taxon:3847"  
/clone="GENOME SYSTEMS CLONE ID: Gm-cl032-3450"  
/clone\_1lb="Gm-cl032"  
/tissue\_type="Cotyledons of 8-day-old 'Williams' seedlings"  
/lab\_host="DH10B"

/note="Vector: pBluescript II XR; Site.1: EcoRI; Site.2: XhoI; This cDNA library was constructed from mRNA isolated from cotyledons of 8-day-old 'Williams' seedlings which were propagated on paper towels with distilled water for 3 days (etiolated), then greenhouse grown for 5 days in potting soil. The cotyledons were flash-frozen in liquid nitrogen. StrataGene's cDNA Synthesis Kit (catalog number 200401) was used to synthesize the cDNA. First-strand synthesis was performed with 5-methyl dCTP, hence the ligated cDNA was hemimethylated. A modification of StrataGene's first-strand synthesis primer was used. An 'anchor' nucleotide (V-A, C or G) was added to the 3' end of the primer [GAGACAGACAGACAGACAGACTGTCGAG(T)18V] to anchor the primer at the 5' end of the poly(A) tract. After second strand synthesis, the cDNA ends were filled in with cloned Pfu DNA, ligated to EcoRI adapters and subsequently phosphorylated. The cDNA was then precipitated and redissolved in sterile, RNase-free water. The XhoI site within the first-strand synthesis primer was then restricted by digestion with XhoI from Promega (400U/ul); all XhoI sites in the cDNA would be protected by their hemimethylated status. The cDNA constructs were size-fractionated with a 500 bp cutoff, using GIBCOBRL Life Technologies' cDNA Size

Fractionation column. The column eluent was then precipitated, redissolved, and ligated into StrataGene's pBluescript II XR predigested vector (pBluescript II SK(+)) vector that has been digested with EcoRI and XhoI, and phosphorylated by StrataGene). 100% of the white and 87.5% of the blue colonies appear to contain recombinant plasmids with cDNA inserts, based on size (n=28 and 8 respectively). This library was constructed by Dr. Paul Keim and Dr. Virginia Corvett.

BASE COUNT 97 a 157 c 120 g 78 t 1 others  
ORIGIN

Query Match 20.2%; Score 42; DB 12; Length 453;  
Best Local Similarity 62.3%; Pred. No. 9.9;  
Matches 66; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

OY 101 CGAGAACAGCTGCTGGCCACATGAGTACTGTCGCTGACGAGGCGCAGACAC 160  
DB 282 CGACACGACGCGCTGTCACGCGCCAGACCTGACGCTCTCAGCTGCTTGGCC 341  
OY 161 GATGACCTGCTGCGCGCGGCGTGCAGTGCCTGTCAGCTGAGACAC 206  
DB 342 GTCCGCGTGGCCGACGACGCTGCTGCTGCGGAGGTGAC 387

RESULT 15  
LOCUS BM326079 468 bp mRNA linear EST 04-JAN-2002

DEFINITION P1C1\_65\_E06.b1 A002 Pathogen-Infected compatible 1 (P1C1) Sorghum

ACCESSION BM326079  
VERSION BM326079.1 GI:18065216

KEYWORDS EST.  
ORGANISM sorghum.

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Sorghum.

REFERENCE 1 (bases 1 to 468)  
AUTHORS Cordonnier-Pratt, M.-M., Gingle, A., Fang, G.C., Dean, R., Wing, R., Sudman, M. and Pratt, L.H.

TITLE An EST database from Sorghum: plants infected with a compatible pathogen  
JOURNAL Unpublished (2002)  
COMMENT Contact: Cordonnier-Pratt MM  
Laboratory for Genomics and Bioinformatics  
The University of Georgia, Department of Plant Biology  
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
Tel: 706 542 1860  
Fax: 706 583 0210  
Email: mmpratt@uga.edu

Sequences have been trimmed to exclude polyA, vector, and regions below phred quality 16. The threshold for highest quality sequence is 20. Three-prime sequences, which are obtained with PolyTmix or T7 sequencing primer, are presented as the reverse complement.  
Seq primer: JEN REV  
High quality sequence stop: 416  
POLYA=NO.

## FEATURES

Location/Qualifiers

1..468  
/organism="Sorghum bicolor"  
/cultivar="Brx623"  
/db\_xref="taxon:4558"  
/clone\_1lb="Pathogen-Infected compatible 1 (P1C1)"  
/tissue\_type="Leaves"  
/dev\_stage="4-week-old seedlings infected with Colletotrichum graminicola"

/note="Vector: pBluescript II SK(-) from Lambda Zap II; Site.1: XhoI; Site.2: EcoRI; Four-week-old sorghum seedlings were sprayed with spore suspension prepared from 3-week-old FRM421, a sorghum isolate of the anthracnose pathogen Colletotrichum graminicola. Inoculated plants were kept in a 25 C dark growth chamber with 100% relative

humidity for 24 hr, followed by 12/12 hr of light/dark cycle at 25 C with 90% relative humidity for another 24 hr. All leaves were harvested and quick frozen with liquid nitrogen and stored in a -80 C freezer. The library was made from poly-A RNA in the cloning vector lambda ZAP II. Clones to be sequenced were prepared by mass excision. WARNING: While most or all ESTs are expected to derive from the host plant, no effort was made to eliminate ESTs deriving from the pathogen.

BASE COUNT 111 a 150 c 136 g 71 t  
ORIGIN

Query Match 20.2%: Score 42; DB 13; Length 468;  
Best Local Similarity 51.6%: Pred. No. 9.9;  
Matches 96: Conservative 0; Mismatches 90; Indels 0; Gaps 0;

OY 22 TGGCCCGCGTGGGCGCCCTACAGGTCACAGAGCTGGGCTGAACGCCGCCAGCCGA 81  
DB 162 TGGACTACTTCACTACGACTACGAGTTCCGCGAGCCCGCGCGTGCAGAGCCTGCAGA 221  
OY 82 TCACGTGTCGACTGTGACCGAGAGAGGTGTGCGCCACCATGAGTACTGTGCGCC 141  
DB 222 ACACGCGAGCCCGAGCGCCACCAAGCGGACTTGGGAGAGACACTACTTCGTGCGGACC 281  
OY 142 TGCACGAGGGCCAGACACGATGACCGTCCCGGCGGCGGTGAGAGTCCCGGTGAGGTGG 201  
DB 282 ACGGTGGTTGAGTGCATCATCCACGCCATCGGCGAGCTCTACTCTCCACGACGCCA 341  
OY 202 ACGACA 207  
DB 342 ACGGCA 347

Search completed: November 13, 2002, 04:00:27  
Job time : 1122.63 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 12, 2002, 16:05:40 : Search time 21.1607 Seconds  
(without alignments)  
3487.380 Million cell updates/sec

Title: US-09-697-123b-19

Sequence: 1 tcaaggagaagcgcctacgac.....ccgctcaggtgacgacat 208

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 320260 seqs, 177392727 residues

Total number of hits satisfying chosen parameters: 640520

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

Published\_Applications\_NA:\*

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*
- 12: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*
- 13: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*
- 14: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	88	42.3	5096	10	US-09-984-711-5
2	88	42.3	5099	9	US-10-075-460-5
3	88	42.3	5099	10	US-09-887-052-1
4	88	42.3	5099	10	US-09-887-052-1
5	88	42.3	5099	10	US-09-887-052-1
6	42.2	20.3	1057	10	US-09-452-239-5
7	42	20.2	2693	10	US-09-880-107-3678
8	41.6	20.0	492	10	US-09-815-242-7946
9	40.4	19.4	392	10	US-09-815-242-7946
10	38.6	18.6	1050	12	US-10-060-857-7
11	38	18.3	1131	10	US-09-934-778-1
12	38	18.3	1467	10	US-09-887-576-777
13	37.8	18.2	484	10	US-09-854-731-1
14	37.8	18.2	1154	10	US-09-815-242-7696
15	37.4	18.0	1535	12	US-10-052-586-545
16	37	17.8	569	10	US-09-864-761-15816
17	37	17.8	1640	10	US-09-925-300-229
18	36.8	17.7	471	10	US-09-815-242-7717
19	36.8	17.7	1266	10	US-09-815-242-7920

20	36.4	17.5	278	10	US-09-923-876-2599	Sequence 2599, Ap
21	36.4	17.5	714	10	US-09-861-289-17	Sequence 17, Appl
22	36.4	17.5	13613	10	US-09-861-289-3	Sequence 3, Appl1
23	36.2	17.4	1350	10	US-09-815-242-7965	Sequence 7965, Ap
24	36	17.3	4074	10	US-09-815-242-7899	Sequence 7899, Ap
25	35.6	17.1	267	10	US-09-923-876-1232	Sequence 1232, Ap
26	35.6	17.1	1162	12	US-10-037-450-21	Sequence 21, Appl
27	35.6	17.1	2010	12	US-10-032-717-9	Sequence 9, Appl1
28	35.4	17.0	1029	10	US-09-815-242-8006	Sequence 8006, Ap
29	35.4	17.0	1822	8	US-08-822-186-1	Sequence 1, Appl1
30	35.4	17.0	1822	8	US-08-937-755-1	Sequence 1, Appl1
31	35.4	17.0	1822	9	US-09-982-543A-9	Sequence 9, Appl1
32	35.4	17.0	1822	10	US-09-045-331-1	Sequence 1, Appl1
33	35.4	17.0	1822	10	US-09-828-607-1	Sequence 1824, Ap
34	35.4	17.0	1822	10	US-09-954-456-1824	Sequence 4086, Ap
35	35.2	16.9	1359	10	US-09-815-242-4086	Sequence 5, Appl1
36	35.2	16.9	2825	10	US-09-952-677-5	Sequence 4952, Ap
37	35	16.8	274	10	US-09-923-876-4952	Sequence 7971, Ap
38	35	16.8	828	10	US-09-815-242-7971	Sequence 3856, Ap
39	34.8	16.7	409	10	US-09-864-761-32324	Sequence 7935, Ap
40	34.8	16.7	1665	10	US-09-815-242-7995	Sequence 109, App
41	34.8	16.7	704	12	US-10-062-254-109	Sequence 111, App
42	34.6	16.6	752	12	US-10-062-254-111	Sequence 7738, Ap
43	34.6	16.6	1437	10	US-09-815-242-7738	Sequence 2873, Ap
44	34.6	16.6				
45	34.4	16.5	261	10	US-09-923-876-2873	

#### ALIGNMENTS

RESULT 1  
US-09-984-711-5  
Sequence 5, Application US/09984711  
Patent No. US20020119549A1

GENERAL INFORMATION:  
APPLICANT: MOECKEL, Bettina  
APPLICANT: BATHE, Brigitte  
APPLICANT: STEPHAN, Hans  
APPLICANT: KREUTER, Caroline  
APPLICANT: HERMANN, Thomas  
APPLICANT: PFEFFERLE, Walter  
APPLICANT: BINDER, Michael  
TITLE OF INVENTION: NUCLEOTIDE SEQUENCES WHICH CODE FOR THE rpsL GENE  
FILE REFERENCE: 204209USO  
CURRENT APPLICATION NUMBER: US/09/984/711  
CURRENT FILING DATE: 2001-10-31  
PRIOR APPLICATION NUMBER: DE10108230.9  
PRIOR FILING DATE: 2001-02-16  
NUMBER OF SEQ ID NOS: 6  
SOFTWARE: Patentin version 3.1  
SEQ ID NO 5  
LENGTH: 5096  
TYPE: DNA  
ORGANISM: Corynebacterium glutamicum  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (702)..(4196)  
OTHER INFORMATION:  
US-09-984-711-5

Query Match  
Best local Similarity 68.5%: Pred. No. 8.4e-14; Length 5096.

Matches 137; Conservative 0; Mismatches 60; Indels 3; Gaps 1;

QY 9 AAGCGTAGACCTGGCCGCGGCGCTACAGGTCACAGAGCTGGCGCTGAAC 68  
DB 1578 AAGCGTAGACCTGGCGCGGCGCTGGCTTACAAAGATCAACCCAGAGCTGGCGCTTGT 1637  
QY 69 GCGGCGACCGCATCATCTGTCGACATCTGACGAGGAAGAGCTGTCGCACCATCGAG 128  
DB 1638 GCGGCGACCGATGCTTATGATGACTCTTACTGAAAGAGCATCGGACCATCGAG 1694





156 GCCACACAGTTCGTCGCCATCAAGGCGCAGGCGGCCAG 115

11111

.....CCTCCACATAGGGGCAAGGGCGGCCGAC 115



```

      b                y                y                b
      |                |                |                |
LOCATION: (389)         NAME/KEY: unsure
NAME/KEY: (396)       LOCATION: (396)
NAME/KEY: unsure     LOCATION: (402)
NAME/KEY: unsure     LOCATION: (430)
NAME/KEY: unsure     LOCATION: (467)
NAME/KEY: unsure     LOCATION: (469)
NAME/KEY: unsure     LOCATION: (479)
US-09-854-731-1

Query Match
Best Local Similarity    18.2%; Score 37.8; DB 10; Length 484;
Matches 95; Conservative 0; Mismatches 77; Indels 1; Gaps 1;

      b                y                y                b
      |                |                |                |
28 GCCTGGGCGCCGTCAAGTCATCACAAGAAGCTGGGCCCTGAACGCCGGCCAACCGATTACGT 87
   |||||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
200 GCGTGGAGCGGGCCACGCAGAGCATTTCCGCTGCAAGTCCATCTCGAAGCGCAACTCGT 259
   |||||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
88 CGTGACCTTGAAACGAGAAAGACGTCTGCGCACCATCGATCGATCTACCTGGTGGTGGCTTCACG 147
   |||||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
260 CACCGACGACGACGA-TGTGGAGGACGTCGCCGCCGGAGATTCACGATTAATGACACCACTCGGGCG 318
   |||||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||

```

Db 319 GCCACCCCAAGCTGATCTCATCCGCGCGCTACAGAGAGCGCCGTCGCCG 371

## RESULT 14

US-09-815-242-7696

Sequence 7696, Application US/09815242  
Patent No. US20020061569A1  
GENERAL INFORMATION:  
APPLICANT: Haselbeck, Robert  
APPLICANT: Ohlson, Karl L.  
APPLICANT: Zyskind, Judith W.  
APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John D.  
APPLICANT: Carr, Grant J.  
APPLICANT: Yamamoto, Robert T.  
APPLICANT: Xu, H. Howard  
TITLE OF INVENTION: Identification of Essential Genes in  
FILE REFERENCE: Prokaryotes  
CURRENT APPLICATION NUMBER: US/09/815,242  
CURRENT FILING DATE: 2001-03-21  
PRIOR APPLICATION NUMBER: 60/191,078  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: 60/206,848  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/207,727  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 60/242,578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253,625  
PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: 60/257,931  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/269,308  
PRIOR FILING DATE: 2001-02-16  
NUMBER OF SEQ ID NOS: 14110  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO: 7696  
LENGTH: 1164  
TYPE: DNA  
ORGANISM: Pseudomonas aeruginosa  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)...(1164)  
US-09-815-242-7696

Query Match Best Local Similarity 18.28; Score 37.8; DB 10; Length 1164;  
Matches 93; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

Db 23 GGGCCGCGTGGCGCTACAGAGTCAAGAGCTGGGCTGAACGCCGCCGCGAT 82  
Db 558 GCGCATGTCGCCGCGCTCAAGGTTTCCACCAAGCTGAGAGCTGTAACAGCTGCCG 617  
Db 83 CACGTCGTCGACTGTACAGGAGAGAGCTGCGCCACCATGAGTACCTGTCGCGCT 142  
Db 618 CATCTGACCAACATGATCTCGCGCGGCGCATGCCAACCTTCTCGCGCGCGG 677  
Db 143 GCAGAGAGGCGAGAGCGAGTACGCGCGCGCGCGCGCGCGCGCGCGCGCG 202  
Db 678 ACACAAAGCTGCGCAAGTCACTGACAGAGCGCGAGCTGTCGAGACCGCCAGCCATCGC 737  
Db 203 CGACA 207  
Db 738 CGCCA 742

## RESULT 15

US-10-052-586-545

Sequence 545, Application US/10052586  
Patent No. US20020127584A1  
GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.  
APPLICANT: Chen, Jian  
APPLICANT: Desnoyers, Luc  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Pan, James  
APPLICANT: Smith, Victoria  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Wood, William I.  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE REFERENCE: P3430R1C1  
CURRENT APPLICATION NUMBER: US/10/052,586  
CURRENT FILING DATE: 2002-01-15  
PRIOR APPLICATION NUMBER: 60/059263  
PRIOR FILING DATE: 1997-09-18  
PRIOR APPLICATION NUMBER: 60/059266  
PRIOR FILING DATE: 1997-09-18  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/063120  
PRIOR FILING DATE: 1997-10-24  
PRIOR APPLICATION NUMBER: 60/063121  
PRIOR FILING DATE: 1997-10-24  
PRIOR APPLICATION NUMBER: 60/063486  
PRIOR FILING DATE: 1997-10-21  
PRIOR APPLICATION NUMBER: 60/063540  
PRIOR FILING DATE: 1997-10-28  
PRIOR APPLICATION NUMBER: 60/063541  
PRIOR FILING DATE: 1997-10-28  
PRIOR APPLICATION NUMBER: 60/063544  
PRIOR FILING DATE: 1997-10-28  
PRIOR APPLICATION NUMBER: 60/063564  
PRIOR FILING DATE: 1997-10-28  
PRIOR APPLICATION NUMBER: 60/063734  
PRIOR FILING DATE: 1997-10-29  
PRIOR APPLICATION NUMBER: 60/063870  
PRIOR FILING DATE: 1997-10-31  
PRIOR APPLICATION NUMBER: 60/064103  
PRIOR FILING DATE: 1997-10-31  
PRIOR APPLICATION NUMBER: 60/065311  
PRIOR FILING DATE: 1997-11-13  
PRIOR APPLICATION NUMBER: 60/066120  
PRIOR FILING DATE: 1997-11-21  
PRIOR APPLICATION NUMBER: 60/064466  
PRIOR FILING DATE: 1997-11-24  
PRIOR APPLICATION NUMBER: 60/066772  
PRIOR FILING DATE: 1997-11-24  
PRIOR APPLICATION NUMBER: 60/069335  
PRIOR FILING DATE: 1997-12-11  
PRIOR APPLICATION NUMBER: 60/069425  
PRIOR FILING DATE: 1997-12-12  
PRIOR APPLICATION NUMBER: 60/069870  
PRIOR FILING DATE: 1997-12-17  
PRIOR APPLICATION NUMBER: 60/068017  
PRIOR FILING DATE: 1997-12-18  
PRIOR APPLICATION NUMBER: 60/077450  
PRIOR FILING DATE: 1998-03-10  
PRIOR APPLICATION NUMBER: 60/077632  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077649  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/078886  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078939  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/079664  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079786  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/080107

;; PRIOR FILING DATE: 1998-03-31  
;; PRIOR APPLICATION NUMBER: 60/080194  
;; PRIOR FILING DATE: 1998-03-31  
;; PRIOR APPLICATION NUMBER: 60/080327  
;; PRIOR FILING DATE: 1998-04-01  
;; PRIOR APPLICATION NUMBER: 60/080333  
;; PRIOR FILING DATE: 1998-04-01  
;; PRIOR APPLICATION NUMBER: 60/081049  
;; PRIOR FILING DATE: 1998-04-08  
;; PRIOR APPLICATION NUMBER: 60/081070  
;; PRIOR FILING DATE: 1998-04-08  
;; PRIOR APPLICATION NUMBER: 60/081195  
;; PRIOR FILING DATE: 1998-04-09  
;; PRIOR APPLICATION NUMBER: 60/081838  
;; PRIOR FILING DATE: 1998-04-15  
;; PRIOR APPLICATION NUMBER: 60/082568  
;; PRIOR FILING DATE: 1998-04-21  
;; PRIOR APPLICATION NUMBER: 60/082569  
;; PRIOR FILING DATE: 1998-04-21  
;; PRIOR APPLICATION NUMBER: 60/082704  
;; PRIOR FILING DATE: 1998-04-22  
;; PRIOR APPLICATION NUMBER: 60/082797  
;; PRIOR FILING DATE: 1998-04-22  
;; PRIOR APPLICATION NUMBER: 60/083322  
;; PRIOR FILING DATE: 1998-04-28  
;; PRIOR APPLICATION NUMBER: 60/083495  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083496  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083499  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083559  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/084366  
;; PRIOR FILING DATE: 1998-05-05  
;; PRIOR APPLICATION NUMBER: 60/084414  
;; PRIOR FILING DATE: 1998-05-06  
;; PRIOR APPLICATION NUMBER: 60/084639  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084640  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084643  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/085573  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085579  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085580  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085582  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085700  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/086023  
;; PRIOR FILING DATE: 1998-05-18  
;; PRIOR APPLICATION NUMBER: 60/086392  
;; PRIOR FILING DATE: 1998-05-22  
;; PRIOR APPLICATION NUMBER: 60/086486  
;; PRIOR FILING DATE: 1998-05-22  
;; PRIOR APPLICATION NUMBER: 60/087098  
;; PRIOR FILING DATE: 1998-05-28  
;; PRIOR APPLICATION NUMBER: 60/087208  
;; PRIOR FILING DATE: 1998-05-28  
;; PRIOR APPLICATION NUMBER: 60/087609  
;; PRIOR FILING DATE: 1998-06-02  
;; PRIOR APPLICATION NUMBER: 60/087759  
;; PRIOR FILING DATE: 1998-06-02  
;; PRIOR APPLICATION NUMBER: 60/087827  
;; PRIOR FILING DATE: 1998-06-03  
;; PRIOR APPLICATION NUMBER: 60/088025  
;; PRIOR FILING DATE: 1998-06-04  
;; PRIOR APPLICATION NUMBER: 60/088028  
;; PRIOR FILING DATE: 1998-06-04

;; PRIOR APPLICATION NUMBER: 60/088029  
;; PRIOR FILING DATE: 1998-06-04  
;; PRIOR APPLICATION NUMBER: 60/088033  
;; PRIOR FILING DATE: 1998-06-04  
;; PRIOR APPLICATION NUMBER: 60/088167  
;; PRIOR FILING DATE: 1998-06-05  
;; PRIOR APPLICATION NUMBER: 60/088202  
;; PRIOR FILING DATE: 1998-06-05  
;; PRIOR APPLICATION NUMBER: 60/088212  
;; PRIOR FILING DATE: 1998-06-05  
;; PRIOR APPLICATION NUMBER: 60/088217  
;; PRIOR FILING DATE: 1998-06-05  
;; PRIOR APPLICATION NUMBER: 60/088326  
;; PRIOR FILING DATE: 1998-06-04  
;; PRIOR APPLICATION NUMBER: 60/088655  
;; PRIOR FILING DATE: 1998-06-09  
;; PRIOR APPLICATION NUMBER: 60/088722  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088738  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088740  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088811  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088824  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088825  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088826  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088861  
;; PRIOR FILING DATE: 1998-06-11  
;; PRIOR APPLICATION NUMBER: 60/088863  
;; PRIOR FILING DATE: 1998-06-11  
;; PRIOR APPLICATION NUMBER: 60/088876  
;; PRIOR FILING DATE: 1998-06-11  
;; PRIOR APPLICATION NUMBER: 60/089090  
;; PRIOR FILING DATE: 1998-06-12  
;; PRIOR APPLICATION NUMBER: 60/089105  
;; PRIOR FILING DATE: 1998-06-12  
;; PRIOR APPLICATION NUMBER: 60/089512  
;; PRIOR FILING DATE: 1998-06-16  
;; PRIOR APPLICATION NUMBER: 60/089514  
;; PRIOR FILING DATE: 1998-06-16  
;; PRIOR APPLICATION NUMBER: 60/089538  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089598  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089653  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089908

Query Match 18.0%; Score 37.4; DB 12; Length 1535;  
Best Local Similarity 53.0%; Pred. No 0.31; Indels 0; Gaps 0;  
Matches 80; Conservative 0; Mismatches 71;

Oy 1 TCACGAGAACGCTACGACTGCGCGCGCTGCGCGCTACAGGTCACAGAACGCTG 60  
Db 852 TCCTGAGAACGCGCTACGCGCGCGCGGAGAGCGCTGAGAGGACAGAGCTATAG 911  
Oy 61 GCGTGAACGCGCGCGAGCGCGATGCGTGCAGTCTGACCGAGAGAGCTGCGCA 120  
Db 912 GCATGATGAGCGCATGAGTACATCAACAGACTGTTGCGCATCGCTCGTCA 971  
Oy 121 CCATGAGTACCTGCTGCGCGCTGCGAGAGG 151  
Db 972 CCATGAGGAGCTGCTGAGATCCACAGGCG 1002

Search completed: November 12, 2002, 16:58:42  
Job time : 33.1607 secs

1	166.4	80.0	970.1	US-08-250-030-1	Sequence 1, Appl1
2	166.4	80.0	970.1	PCT-US95-06790-1	Sequence 1, Appl1
3	142.4	68.5	3447.2	US-08-313-185-57	Sequence 57, Appl1
4	142.4	68.5	3447.2	US-09-082-614A-57	Sequence 57, Appl1
5	42.8	20.3	1734.6	5352575-8	Patent No. 5352575
6	42.2	20.3	1057.7	US-09-450-239-5	Sequence 5, Appl1
7	42.2	20.3	4411529.4	US-09-193-940A-1	Sequence 1, Appl1
8	38.8	18.7	1450.1	US-07-923-652C-5	Sequence 5, Appl1
9	38.8	18.7	1450.1	US-08-184-237-5	Sequence 5, Appl1
10	38.8	18.7	1450.1	US-08-482-920-5	Sequence 5, Appl1
11	38.8	18.7	1450.3	US-08-484-341-5	Sequence 5, Appl1
12	38.8	18.7	2742.4	US-09-233-468A-1	Sequence 1, Appl1
13	38.8	18.7	4897.6	5196516-7	Patent No. 5196516
14	38.8	18.7	6085.4	US-09-028-603-4	Sequence 4, Appl1
15	38.6	18.6	1164.1	US-07-640-476-6	Sequence 6, Appl1
16	38.6	18.3	1131.4	US-09-420-211-1	Sequence 1, Appl1
17	37.8	18.2	484.4	US-09-347-801-1	Sequence 1, Appl1
18	37.4	18.0	971.3	US-09-248-335-65	Sequence 65, Appl1
19	37.2	17.9	1452.4	US-08-483-502-5	Sequence 5, Appl1
20	37.2	17.9	1452.4	US-08-726-651A-5	Sequence 5, Appl1
21	37.2	17.9	1452.4	US-09-079-955-10	Sequence 10, Appl1
22	37.2	17.9	3000.4	US-08-192-104-1	Sequence 1, Appl1
23	37.2	17.9	3000.4	US-09-543-446-1	Sequence 1, Appl1
24	37.2	17.9	4276.1	US-07-973-324A-3	Sequence 3, Appl1
25	37.2	17.9	4276.4	US-08-343-380-3	Sequence 3, Appl1
26	37.2	17.9	4276.4	US-09-072-435-3	Sequence 3, Appl1
27	37.2	17.9	4276.4	US-09-072-917A-3	Sequence 3, Appl1

[illegible]

Query Match	80.0%	Score 166.4	DB 1	Length 970
Best Local Similarity	87.5%	Pred. No. 5.3e-31		
Matches 182	Conservative	0	Mismatches 26	Indels 0
			Gaps	0
OY	1	TCAGGGAAGAGCCCTACGACCTGCGCCCGCGTGAGCCGCTACAGTCACAAAGAACTGG	60	
Db	26	TCAGGGAAGAGCCCTACGACCTGCGCCCGCGTGCGCTATTAAGCTCAACAAAGACTTTC	85	









Query Match	18.7%	Score 38.8;	DB 1;	Length 1450;
Best Local Similarity	52.5%;	Pred. No. 0.45;		
Matches 85;	Conservative	0;	Mismatches 77;	Indels 0;
			Gaps	0
QY 35	CCGCTACAGGTCAACAAGACGCTGGCTGAAGCCCGCCAGCCATCACGTCGTGAC	94		
Db 891	CCGGGGCGAGAGATGCGAAGGGCCCGGCATATCGGGTGTGGCGCCGCGCAAGCGCGAC	950		
QY 95	TCGTACACGAGGAAGAGCTGTCGCCACCATCGAGTACGTGTGGCTGTGACGAGGCCA	154		
Db 951	CTTGTGTGACAAACGACGACACCGGCTGCACGCACGACACTGTGGCCCTTCCCTCCACAA	1010		
QY 155	GACCACAGTAGACCGTCCCGCGCGCGGTGAGGTCCCGGTGCA	196		
Db 1011	GGTCATATGCAAGGGCTTACGATATCATCTCATCCACCCGGCA	1052		

```

? FILLING DATE: <Unknown> US 600,244
? APPLICATION NUMBER: US 641,617
? FILLING DATE: 22-OCT-1990
? APPLICATION NUMBER: US 310,881
? FILLING DATE: 16-JAN-1991
? APPLICATION NUMBER: US 160,766
? FILLING DATE: 17-FEB-1989
? APPLICATION NUMBER: US 160,771
? FILLING DATE: 26-FEB-1988
? APPLICATION NUMBER: US 347,637
? FILLING DATE: 05-MAY-1989
? APPLICATION NUMBER: US 363,138
? FILLING DATE: 08-JUN-1989
? APPLICATION NUMBER: US 219,279
? FILLING DATE: 15-JUL-1988
? ATTORNEY/AGENT INFORMATION:
? NAME: Halliwin, Albert P.
? REGISTRATION NUMBER: 28,957
? REFERENCE/DOCKET NUMBER: BIOG-20121 USA
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 415-433-4150
? TELEFAX: 415-433-8716
? INFORMATION FOR SEQ ID NO: 5:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1450 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: DNA (genomic)
? HYPOTHEetical: NO
? ANTI-SENSE: NO
? ORIGINAL SOURCE:
? ORGANISM: Oryza sativa
? IMMEDIATE SOURCE:
? CLONE: alpha-amy1ase
? FEATURE:
? NAME/KEY: CDS (B) LOCATION: 12..1316
? LOCATION: 12..1316
? SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-08-484-341-5

Query Match 18.7% Score 38.8; DB 3; Length 1450;
Best Local similarity 52.5% Pred. No. 0.45;
Matches 85; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

QY 35 CCGTCAAGGTCACAAAGAGCTGGGCGCTGAAAGCCGGCCAGCGCATACGTCGTGCAC 94
DB 891 CCGCGCGGAGGACGCGGAGCGCGCGGCGCATATCGGGTCTGGCGCGCCAAAGCGAGCGAC 950
QY 95 TCTGACGAGGAGAAAGACTGCTGCCACCAATCGAGTACCTGTGTGGCGCTGCACGAGGAGCA 154
DB 951 CTGTGTCGACACACGACGACACCGGCTCGACGCGACGACCTGTGGCGCGTTCCCTCGACAA 1010
QY 155 GACCACGATGACCGTCCCGCGGCGGCTCGAGAGTCCCGGTGCA 196
DB 1011 GGTGATGACAGGCTACGCAATACATCTCTACACCCAGCCCGGCA 1052

RESULT 12
US-09-232-468A-1
? Sequence 1, Application US/09232468A
? Patent No. 6207165
? GENERAL INFORMATION:
? APPLICANT: AUDONNET et al.
? TITLE OF INVENTION: POLYNUCLEOTIDE VACCINE FORMULA AGAINST PORCINE
? FILE REFERENCE: 454313-2230
? CURRENT APPLICATION NUMBER: US/09/232,468A
? NUMBER OF SEQ ID NOS: 54
? SOFTWARE: PatentIn Ver. 2.1

```





PT diagnosis and identification of many mycobacterial species by restriction fragment length polymorphism -

XX Claim 1; Page 46; 50pp; English.

PS The present sequence for *Mycobacterium fortuitum* rpoB gene  
CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from  
CC various *Mycobacterium* species. These rpoB gene fragments can be used  
CC in the diagnosis and identification of *Mycobacterium* species using a  
CC novel PCR-restriction fragment length polymorphism analysis (PRA)  
CC method. The method comprises obtaining a restriction fragment length  
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,  
CC amplifying and digesting the DNA fragment from the microorganism to  
CC be identified and comparing the RFLP patterns from the known rpoB gene  
CC fragments with the unidentified fragment. The rpoB gene fragments  
CC are useful to identify a wide range of *Mycobacterium* species, e.g. for  
CC diagnosis or to obtain epidemiological and pathogenesis information for  
CC selection of appropriate therapies, including *M. tuberculosis*, *M. leprae*  
CC and non-tuberculous *Mycobacterium* (NTM) encountered in subjects infected  
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene  
CC fragments is rapid, precise, simple and cost effective (only 1 PCR  
CC experiment), and can differentiate between many species in a single  
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for  
CC detecting specific *Mycobacterium* species.

XX Sequence 208 BP; 43 A; 69 C; 69 G; 27 T; 0 other;

Query Match 100.0%; Score 208; DB 22; Length 208;  
Best Local Similarity 100.0%; Pred. No. 1e-37;

Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCAAGAGAACGGCTACGACCTGGCCCGCGCTGACGACGATGACAGAGAGCTGG 60  
DB 1 TCAAGAGAACGGCTACGACCTGGCCCGCGCTGACGACGATGACAGAGAGCTGG 60  
QY 61 GCCTGACGCGCGCCAGCCGATCACTGCTGACGACGAGAGAGCTGGTGGCCA 120  
DB 61 GCCTGACGCGCGCCAGCCGATCACTGCTGACGACGAGAGAGCTGGTGGCCA 120  
QY 121 CCATGAGTACTGCTGGTGGCTGACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180  
DB 121 CCATGAGTACTGCTGGTGGCTGACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180  
QY 181 TCGAGGTCCCGGTGAGGTGAGACAT 208  
DB 181 TCGAGGTCCCGGTGAGGTGAGACAT 208

# RESULT 2

AAS05215 standard; DNA; 208 BP.

AC AAS05215;

DT 07-SEP-2001 (first entry)

DE *Mycobacterium avium* rpoB gene fragment.

KW Non-tuberculous *Mycobacterium*; rpoB gene fragment; NTM; HIV; PRA; RFLP;  
KW PCR-restriction fragment length polymorphism analysis; ds.

OS *Mycobacterium avium*.

PN WO200131061-A1.

PD 03-MAY-2001.

PF 27-OCT-2000; 2000WO-KR01223.

PR 27-OCT-1999; 99KR-0046795.

PA (ERUM-) ERUM BIOTECH CO LTD.

PI Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;

XX WPI; 2001-300520/31.

DR New DNA fragments from the rpoB gene of *Mycobacterium*, useful for  
XX diagnosis and identification of many *Mycobacterium* species by  
PT restriction fragment length polymorphism.

PS Claim 1; Page 44; 50pp; English.

CC The present sequence for *Mycobacterium avium* rpoB gene  
CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from  
CC various *Mycobacterium* species. These rpoB gene fragments can be used  
CC in the diagnosis and identification of *Mycobacterium* species using a  
CC novel PCR-restriction fragment length polymorphism analysis (PRA)  
CC method. The method comprises obtaining a restriction fragment length  
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,  
CC amplifying and digesting the DNA fragment from the microorganism to  
CC be identified and comparing the RFLP patterns from the known rpoB gene  
CC fragments with the unidentified fragment. The rpoB gene fragments  
CC are useful to identify a wide range of *Mycobacterium* species, e.g. for  
CC diagnosis or to obtain epidemiological and pathogenesis information for  
CC selection of appropriate therapies, including *M. tuberculosis*, *M. leprae*  
CC and non-tuberculous *Mycobacterium* (NTM) encountered in subjects infected  
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene  
CC fragments is rapid, precise, simple and cost effective (only 1 PCR  
CC experiment), and can differentiate between many species in a single  
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for  
CC detecting specific *Mycobacterium* species.

XX Sequence 208 BP; 44 A; 69 C; 69 G; 26 T; 0 other;

Query Match 86.2%; Score 179.2; DB 22; Length 208;  
Best Local Similarity 91.3%; Pred. No. 2.7e-31;

Matches 190; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 1 TCAAGAGAACGGCTACGACCTGGCCCGCGCTGACGACGATGACAGAGAGCTGG 60  
DB 1 TCAAGAGAACGGCTACGACCTGGCCCGCGCTGACGACGATGACAGAGAGCTGG 60  
QY 61 GCCTGACGCGCGCCAGCCGATCACTGCTGACGACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120  
DB 61 GCCTGACGCGCGCCAGCCGATCACTGCTGACGACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120  
QY 121 CCATGAGTACTGCTGGTGGCTGACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180  
DB 121 CCATGAGTACTGCTGGTGGCTGACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180  
QY 181 TCGAGGTCCCGGTGAGGTGAGACAT 208  
DB 181 TCGAGGTCCCGGTGAGGTGAGACAT 208

# RESULT 3

AAS05211 standard; DNA; 208 BP.

AC AAS05211;

DT 07-SEP-2001 (first entry)

DE *Mycobacterium marinum* rpoB gene fragment.

KW Non-tuberculous *Mycobacterium*; rpoB gene fragment; NTM; HIV; PRA; RFLP;  
KW PCR-restriction fragment length polymorphism analysis; ds.

OS *Mycobacterium marinum*.

PN WO200131061-A1.

PD 03-MAY-2001.

PF 27-OCT-2000; 2000WO-KR01223.



XX 27-OCT-1999; 99KR-0046795.  
 XX (ERUM-) ERUME BIOTECH CO LTD.

XX Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;  
 XX WPI: 2001-300520/31.

PT New DNA fragments from the rpoB gene of mycobacteria, useful for  
 PT diagnosis and identification of many mycobacterial species by  
 PT restriction fragment length polymorphism

PS Claim 1; Page 43; 50pp; English.

CC The present sequence for *Mycobacterium marinum* rpoB gene  
 CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05244) from  
 CC various *Mycobacterium* species. These rpoB gene fragments can be used  
 CC in the diagnosis and identification of *Mycobacterium* species using a  
 CC novel PCR-restriction fragment length polymorphism analysis (PRA)  
 CC method. The method comprises obtaining a restriction fragment length  
 CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,  
 CC amplifying and digesting the DNA fragment from the microorganism to  
 CC be identified and comparing the RFLP patterns from the known rpoB gene  
 CC fragments with the unidentified fragment. The rpoB gene fragments  
 CC are useful to identify a wide range of *Mycobacterium* species, e.g. for  
 CC diagnosis or to obtain epidemiological and pathogenesis information for  
 CC selection of appropriate therapies, including *M. tuberculosis*, *M. leprae*  
 CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected  
 CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene  
 CC fragments is rapid, precise, simple and cost effective (only 1 PCR  
 CC required), and can differentiate between many species in a single  
 CC experiment, including those difficult to distinguish by usual biochemical  
 CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for  
 CC detecting specific *Mycobacterium* species.

SO Sequence 208 BP; 45 A; 69 C; 68 G; 26 T; 0 other;

Query Match 83.8%; Score 174.4; DB 22; Length 208;  
 Best Local Similarity 89.9%; Pred. No. 3.1e-30;  
 Matches 187; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

OY 1 TCAAGAGAACGCTACGACCTGGCCCGCGCTTACAGGTCAACAGAAAGCTGG 60  
 DB 1 TCAAGAGAACGCTACGACCTGGCCCGCGGTGGCGGTACAAAGTCAACAGAAAGCTGG 60  
 OY 61 GCCTGAAGCGCGGCGACCGCATCAGCTGCTGACCGAGGAAGACGTCGCGCA 120  
 DB 61 GCCTGAAGCGCGGCGACCGCATCAGCTGCTGACCGAGGAAGACGTCGCGCA 120  
 OY 121 CCATCGAGTACCTGTCGCGCTGCAGAGAGGCGACGACGATGACCGTCCCGCGGCGG 180  
 DB 121 CCATCGAATACCTGTCGCGCTGCAGAGAGGCGGCGACGCGATGACCGTCCCGCGGCGG 180  
 OY 181 TCGAGGTCCGCGTTCGAGGTGAGAGACAT 208  
 DB 181 TCGAGGTCCGCGTTCGAGGTGAGAGACAT 208

RESULT 4  
 AAS05210  
 ID AAS05210 standard; DNA: 208 BP.

XX AAS05210;

DT 07-SEP-2001 (first entry)

DE *Mycobacterium ulcerans* rpoB gene fragment.

XX Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;  
 KM PCR-restriction fragment length polymorphism analysis; ds.  
 OS *Mycobacterium ulcerans*.

XX MO200131061-A1.  
 PN 03-MAY-2001.

XX 27-OCT-2000; 2000WO-KR01223.

XX 27-OCT-1999; 99KR-0046795.

XX (ERUM-) ERUME BIOTECH CO LTD.

PI Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;  
 XX WPI: 2001-300520/31.

PT New DNA fragments from the rpoB gene of mycobacteria, useful for  
 PT diagnosis and identification of many mycobacterial species by  
 PT restriction fragment length polymorphism

PS Claim 1; Page 43; 50pp; English.

CC The present sequence for *Mycobacterium ulcerans* rpoB gene  
 CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05244) from  
 CC various *Mycobacterium* species. These rpoB gene fragments can be used  
 CC in the diagnosis and identification of *Mycobacterium* species using a  
 CC novel PCR-restriction fragment length polymorphism analysis (PRA)  
 CC method. The method comprises obtaining a restriction fragment length  
 CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,  
 CC amplifying and digesting the DNA fragment from the microorganism to  
 CC be identified and comparing the RFLP patterns from the known rpoB gene  
 CC fragments with the unidentified fragment. The rpoB gene fragments  
 CC are useful to identify a wide range of *Mycobacterium* species, e.g. for  
 CC diagnosis or to obtain epidemiological and pathogenesis information for  
 CC selection of appropriate therapies, including *M. tuberculosis*, *M. leprae*  
 CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected  
 CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene  
 CC fragments is rapid, precise, simple and cost effective (only 1 PCR  
 CC required), and can differentiate between many species in a single  
 CC experiment, including those difficult to distinguish by usual biochemical  
 CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for  
 CC detecting specific *Mycobacterium* species.

SO Sequence 208 BP; 45 A; 68 C; 67 G; 28 T; 0 other;

Query Match 83.1%; Score 172.8; DB 22; Length 208;  
 Best Local Similarity 89.4%; Pred. No. 7.1e-30;  
 Matches 186; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

OY 1 TCAAGAGAACGCTACGACCTGGCCCGCGCTTACAGGTCAACAGAAAGCTGG 60  
 DB 1 TCAAGAGAACGCTACGACCTGGCCCGCGGTGGCGGTACAAAGTCAACAGAAAGCTGG 60  
 OY 61 GCCTGAAGCGCGGCGACCGCATCAGCTGCTGACCGAGGAAGACGTCGCGCA 120  
 DB 61 GCCTGAAGCGCGGCGACCGCATCAGCTGCTGACCGAGGAAGACGTCGCGCA 120  
 OY 121 CCATCGAGTACCTGTCGCGCTGCAGAGAGGCGACGACGATGACCGTCCCGCGGCGG 180  
 DB 121 CCATCGAATACCTGTCGCGCTGCAGAGAGGCGGCGACGCGATGACCGTCCCGCGGCGG 180  
 OY 181 TCGAGGTCCGCGTTCGAGGTGAGAGACAT 208  
 DB 181 TCGAGGTCCGCGTTCGAGGTGAGAGACAT 208

RESULT 5  
 AAS05201  
 ID AAS05201 standard; DNA: 208 BP.

XX AAS05201;

DT 07-SEP-2001 (first entry)

DE Mycobacterium gordonae type I rpoB gene fragment.

XX Non-tuberculous mycobacteria: rpoB gene fragment; NTM; HIV; PRA; RFLP;  
KW PCR-restriction fragment length polymorphism analysis; ds.

XX Mycobacterium gordonae type I.

OS WO200131061-A1.

PN 03-MAY-2001.

PD 27-OCT-2000; 2000WO-KR01223.

PR 27-OCT-1999; 99KR-0046795.

XX (ERUM-) ERUME BIOTECH CO LTD.

XX Lee H, Park YK, Bal G, Kim S, Cho S, Kim Y, Park HJ;

DR WPI: 2001-300520/31.

PT New DNA fragments from the rpoB gene of mycobacteria, useful for  
PT diagnosis and identification of many mycobacterial species by  
XX restriction fragment length polymorphism

PS Claim 1: Page 40; 50pp; English.

XX The present sequence for Mycobacterium gordonae type I rpoB gene  
CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from  
CC various Mycobacterial species. These rpoB gene fragments can be used  
CC in the diagnosis and identification of Mycobacterium species using a  
CC novel PCR-restriction fragment length polymorphism analysis (PRA)  
CC method. The method comprises obtaining a restriction fragment length  
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,  
CC amplifying and digesting the DNA fragment from the microorganism to  
CC be identified and comparing the RFLP patterns from the known rpoB gene  
CC fragments with the unidentified fragment. The rpoB gene fragments  
CC are useful to identify a wide range of Mycobacterium species, e.g. for  
CC diagnosis or to obtain epidemiological and pathogenesis information for  
CC selection of appropriate therapies, including M. tuberculosis, M. leprae  
CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected  
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene  
CC fragments is rapid, precise, simple and cost effective (only 1 PCR  
CC required), and can differentiate between many species in a single  
CC experiment, including those difficult to distinguish by usual biochemical  
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for  
CC detecting specific Mycobacterial species.

SO Sequence 208 BP; 46 A; 71 C; 65 G; 26 T; 0 other;

Query Match 81.5%; Score 169.6; DB 22; Length 208;  
Best Local Similarity 88.5%; Pred. No. 3.7e-29;

Matches 184; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

```
OY 1 TCAAGGAGAACCGCTAGACCTGCGCCGCTGGCGCGCTACAGATCAACAGAACTCG 60
DB 1 TCAAGGAGAACCGCTAGACCTGCGCCGCTGGCGCGCTACAGATCAACAGAACTCG 60
OY 61 GCCTGAACGCGCGCAGCGATCGTGCATCGTGCACCGAGGAAAGAGCTGTGCCA 120
DB 61 GCCTGAACGCGCGCAGCGATCGTGCATCGTGCACCGAGGAAAGAGCTGTGCCA 120
OY 121 CCATCGAGTACTGTGTGCGCTGCACGAGGCCAGACGATGACCTGCCGCGGCGG 180
DB 121 CCATCGAGTACTGTGTGCGCTGCACGAGGCCAGACGATGACCTGCCGCGGCGG 180
OY 181 TCGAGGTCCCGTGCAGAGTGCAGACAT 208
DB 181 CCGAGGTCCCGTGCAGAGTGCAGACAT 208
```

RESULT 6  
AAS05202

ID AAS05202 standard; DNA: 208 BP.

AC AAS05202;

DT 07-SEP-2001 (first entry)

DE Mycobacterium gordonae type II rpoB gene fragment.

XX Non-tuberculous mycobacteria: rpoB gene fragment; NTM; HIV; PRA; RFLP;  
KW PCR-restriction fragment length polymorphism analysis; ds.

OS Mycobacterium gordonae type II.

PN WO200131061-A1.

PD 03-MAY-2001.

PF 27-OCT-2000; 2000WO-KR01223.

PR 27-OCT-1999; 99KR-0046795.

XX (ERUM-) ERUME BIOTECH CO LTD.

XX Lee H, Park YK, Bal G, Kim S, Cho S, Kim Y, Park HJ;

DR WPI: 2001-300520/31.

PT New DNA fragments from the rpoB gene of mycobacteria, useful for  
PT diagnosis and identification of many mycobacterial species by  
XX restriction fragment length polymorphism

PS Claim 1: Page 40; 50pp; English.

XX The present sequence for Mycobacterium gordonae type II rpoB gene  
CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from  
CC various Mycobacterial species. These rpoB gene fragments can be used  
CC in the diagnosis and identification of Mycobacterium species using a  
CC novel PCR-restriction fragment length polymorphism analysis (PRA)  
CC method. The method comprises obtaining a restriction fragment length  
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,  
CC amplifying and digesting the DNA fragment from the microorganism to  
CC be identified and comparing the RFLP patterns from the known rpoB gene  
CC fragments with the unidentified fragment. The rpoB gene fragments  
CC are useful to identify a wide range of Mycobacterium species, e.g. for  
CC diagnosis or to obtain epidemiological and pathogenesis information for  
CC selection of appropriate therapies, including M. tuberculosis, M. leprae  
CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected  
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene  
CC fragments is rapid, precise, simple and cost effective (only 1 PCR  
CC required), and can differentiate between many species in a single  
CC experiment, including those difficult to distinguish by usual biochemical  
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for  
CC detecting specific Mycobacterial species.

SO Sequence 208 BP; 45 A; 62 C; 73 G; 28 T; 0 other;

Query Match 81.5%; Score 169.6; DB 22; Length 208;  
Best Local Similarity 88.5%; Pred. No. 3.7e-29;

Matches 184; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

```
OY 1 TCAAGGAGAACCGCTAGACCTGCGCCGCTGGCGCGCTACAGATCAACAGAACTCG 60
DB 1 TCAAGGAGAACCGCTAGACCTGCGCCGCTGGCGCGCTACAGATCAACAGAACTCG 60
OY 61 GCCTGAACGCGCGCAGCGATCGTGCATCGTGCACCGAGGAAAGAGCTGTGCCA 120
DB 61 GCCTGAACGCGCGCAGCGATCGTGCATCGTGCACCGAGGAAAGAGCTGTGCCA 120
OY 121 CCATCGAGTACTGTGTGCGCTGCACGAGGCCAGACGATGACCTGCCGCGGCGG 180
DB 121 CCATCGAGTACTGTGTGCGCTGCACGAGGCCAGACGATGACCTGCCGCGGCGG 180
OY 181 TCGAGGTCCCGTGCAGAGTGCAGACAT 208
```

Db 181 CCGAGGTGCCGTGAGACCGACGACAT 208

## RESULT 7

ID AAS05208 standard; DNA: 208 BP.

AC AAS05208;

DT 07-SEP-2001 (first entry)

DE Mycobacterium kansasii rpoB gene fragment.

KW Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;  
PCR-restriction fragment length polymorphism analysis; ds.

OS Mycobacterium kansasii.

PN WO200131061-A1.

PD 03-MAY-2001.

PF 27-OCT-2000; 2000WO-KR01223.

PR 27-OCT-1999; 99KR-0046795.

PA (ERUM-) ERUME BIOTECH CO LTD.

PI Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;

DR WPI: 2001-300520/31.

PT New DNA fragments from the rpoB gene of mycobacteria, useful for  
diagnosis and identification of many mycobacterial species by  
restriction fragment length polymorphism

PS Claim 1: Page 42; 50pp; English.

CC The present sequence for Mycobacterium kansasii rpoB gene  
fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from  
various Mycobacterial species. These rpoB gene fragments can be used  
in the diagnosis and identification of Mycobacterium species using a  
novel PCR-restriction fragment length polymorphism analysis (PRA)  
method. The method comprises obtaining a restriction fragment length  
polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,  
amplifying and digesting the DNA fragment from the microorganism to  
be identified and comparing the RFLP patterns from the known rpoB gene  
fragments with the unidentified fragment. The rpoB gene fragments  
are useful to identify a wide range of Mycobacterium species, e.g. for  
diagnosis or to obtain epidemiological and pathogenesis information for  
selection of appropriate therapies, including M. tuberculosis, M. leprae  
and non-tuberculous mycobacteria (NTM) encountered in subjects infected  
with human immunodeficiency virus (HIV). Analysis of the rpoB gene  
fragments is rapid, precise, simple and cost effective (only 1 PCR  
required), and can differentiate between many species in a single  
experiment, including those difficult to distinguish by usual biochemical  
tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for  
detecting specific Mycobacterial species.

CC Sequence 208 BP; 51 A; 65 C; 65 G; 27 T; 0 other;

Query Match 80.8%; Score 168; DB 22; Length 208;  
Best Local Similarity 88.0%; Pred. No. 8.3e-29;

Matches 183; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

OY 1 TCAAGAGAGGCTACGACCTGCGCCGCGCTACAGGTCAACAGAGAGCTGG 60

DB 1 TCAAGAGAGGCTACGACCTGCGCCGCGCTACAGGTCAACAGAGAGCTGG 60

OY 61 GCTTACAGCGCGCCGACGATCATGCTGACTGACCGAGAGAGAGCTGCGCCA 120

DB 61 GCTTACAGCGCGCGACATCATGCTGACGACGAGAGAGAGAGCTGCGCCA 120

OY 121 CCATGAGTACTGTCGCGCTGCACGAGGCCAGACGATGACCTGCCCGCGCGC 180  
DB 121 CCATCGAGTATCTGTCGCGCTGCACGAGGCCAGACGATGACCTGCCCGCGCGC 180  
OY 181 TCGAGGTCCGCGTGCAGGTGACGACAT 208  
DB 181 TCGAGGTCCGCGTGCAGAACCGACGACAT 208

## RESULT 8

ID AAS05217 standard; DNA: 208 BP.

AC AAS05217;

DT 07-SEP-2001 (first entry)

DE Mycobacterium celatum rpoB gene fragment.

KW Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;  
PCR-restriction fragment length polymorphism analysis; ds.

OS Mycobacterium celatum.

PN WO200131061-A1.

PD 03-MAY-2001.

PF 27-OCT-2000; 2000WO-KR01223.

PR 27-OCT-1999; 99KR-0046795.

PA (ERUM-) ERUME BIOTECH CO LTD.

PI Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;

DR WPI: 2001-300520/31.

PT New DNA fragments from the rpoB gene of mycobacteria, useful for  
diagnosis and identification of many mycobacterial species by  
restriction fragment length polymorphism

PS Claim 1: Page 45; 50pp; English.

CC The present sequence for Mycobacterium celatum rpoB gene  
fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from  
various Mycobacterial species. These rpoB gene fragments can be used  
in the diagnosis and identification of Mycobacterium species using a  
novel PCR-restriction fragment length polymorphism analysis (PRA)  
method. The method comprises obtaining a restriction fragment length  
polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,  
amplifying and digesting the DNA fragment from the microorganism to  
be identified and comparing the RFLP patterns from the known rpoB gene  
fragments with the unidentified fragment. The rpoB gene fragments  
are useful to identify a wide range of Mycobacterium species, e.g. for  
diagnosis or to obtain epidemiological and pathogenesis information for  
selection of appropriate therapies, including M. tuberculosis, M. leprae  
and non-tuberculous mycobacteria (NTM) encountered in subjects infected  
with human immunodeficiency virus (HIV). Analysis of the rpoB gene  
fragments is rapid, precise, simple and cost effective (only 1 PCR  
required), and can differentiate between many species in a single  
experiment, including those difficult to distinguish by usual biochemical  
tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for  
detecting specific Mycobacterial species.

CC Sequence 208 BP; 48 A; 71 C; 64 G; 25 T; 0 other;

Query Match 80.8%; Score 168; DB 22; Length 208;  
Best Local Similarity 88.0%; Pred. No. 8.3e-29;

Matches 183; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

OY 1 TCAAGAGAGGCTACGACCTGCGCCGCGCTACAGGTCAACAGAGAGCTGG 60

Query Match	Similarity	80.0%	Score 166.4	DB 22:	Length 208:
Best Local	Similarity	87.5%	Pred. No. 1.9e-28:		
Matches 182:	Conservative	0:	Mismatch	26:	Indels 0: Gaps 0:
OY	1	TCACGAGAACGCGCTACGACCTGCGCCGCGTACAGAGTTCACACAAAGAACTCG	60		
DB	1	TCACGAGAACGCGCTACGACCTGCGCCGCGTACAGAGTTCACACAAAGAACTCG	60		
OY	61	GGCTTAACGCGCGCCAGCCGATCAGTGTGCACTGTGACCCGAGAAAGCTGTGCGCA	120		
DB	61	GGCTTAACGCGCGCCAGCCGATCAGTGTGCACTGTGACCCGAGAAAGCTGTGCGCA	120		
OY	121	CCATGAGTACCTGTGCTGCGCTGACAGAGGCGCCAGACACGATGACCGTCCCGCGCG	180		
DB	121	CCATGAGTATCTGTGCTGCGCTGACAGAGGCTGACACCGATGACCGTTCGCGCGCG	180		
OY	181	TCGAGGTCGCCGCTGAGAGTGCACGACAT	208		
DB	181	TCGAGGTCGCCGCTGAGAAACGACGACAT	208		
RESULT 10					
AA505206					
ID	AA505206	standard: DNA: 208 BP.			
XX	AA505206:				
AC					
XX					
DT	07-SEP-2001 (first entry)				
XX					
DE	Mycobacterium terrae rpoB gene fragment.				
XX					
KM	Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;				
XX	KCR-restriction fragment length polymorphism analysis; ds.				
OS	Mycobacterium terrae.				
XX					
PN	W0200131061-A1.				
XX					
PD	03-MAY-2001.				
XX					
PF	27-OCT-2000; 2000MO-KR01223.				
XX					
PR	27-OCT-1999; 99KR-0046795.				
XX					
PA	(ERDM-) ERDME BIOTECH CO LTD.				
XX					
PI	Lee H, Park YK, Bal G, Kim S, Cho S, Kim Y, Park HJ.				
XX					
DR	WPI: 2001-300520/31.				
XX					
PT	New DNA fragments from the rpoB gene of mycobacteria, useful for				
PT	diagnosis and identification of many mycobacterial species by				
PT	restriction fragment length polymorphism -				
XX					
PS	Claim 1; Page 42; 50pp; English.				
XX					
CC	The present sequence for Mycobacterium terrae rpoB gene				
CC	fragment is 1 of 24 rpoB gene fragments (AA505201-AA505224) from				
CC	various mycobacterial species. These rpoB gene fragments can be used				
CC	in the diagnosis and identification of Mycobacterium species using a				
CC	novel PCR-restriction fragment length polymorphism analysis (PRA)				
CC	method. The method comprises obtaining a restriction fragment length				
CC	polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,				
CC	amplifying and digesting the DNA fragment from the microorganism to				
CC	be identified and comparing the RFLP patterns from the known rpoB gene				
CC	fragments with the unidentified fragment. The rpoB gene fragments				
CC	are useful to identify a wide range of Mycobacterium species, e.g. for				
CC	diagnosis or to obtain epidemiological and pathogenesis information for				
CC	selection of appropriate therapies, including M. tuberculosis, M. leprae				
CC	and non-tuberculous mycobacteria (NTM) encountered in subjects infected				
CC	with human immunodeficiency virus (HIV). Analysis of the rpoB gene				
CC	fragments is rapid, precise, simple and cost effective (only 1 PCR				

CC required), and can differentiate between many species in a single  
 CC experiment, including those difficult to distinguish by usual biochemical  
 CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for  
 CC detecting specific Mycobacterial species.

SO Sequence 208 BP; 46 A; 62 C; 67 G; 33 T; 0 other;

Query Match 80.0%; Score 166.4; DB 22; Length 208;  
 Best Local Similarity 87.5%; Pred. No. 1.9e-28;  
 Matches 182; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

OY 1 TCAGAGGAAGCGCTACGACCTGGCCCGCGGTGGCGCTACAGGTCAACAGAAAGCTGG 60  
 DB 1 TCAGAGGAAGCGCTACGACCTGGCCCGCGGTGGCGCTACAGGTCAACAGAAAGCTGG 60  
 OY 61 GCCTGAACCGCCGCGCATCAGCTGCTGCTACCTCTACCGAGAGAGAGCTGCGCA 120  
 DB 61 GCCTGATGTGCGCGCATCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120  
 OY 121 CCATCGAGTACCTGCTGCGCTGCGACGAGGCGGACAGCAGTACGCTGCGCGCGCG 180  
 DB 121 CCATCGAATATCTGCTGCGCTGCGACGAGGCGGACAGCAGTACGCTGCGCGCGCG 180  
 OY 181 TCGAGGTCCCGGTGAGGTGAGCAGCAT 208  
 DB 181 TCGAGGTCCCGGTGAGGTGAGCAGCAT 208

RESULT 11  
 AAS05216  
 ID AAS05216 standard; DNA: 208 BP.

AC AAS05216;

DT 07-SEP-2001 (first entry)

DE Mycobacterium bovis rpoB gene fragment.

XX Non-tuberculous mycobacteria: rpoB gene fragment; NTM; HIV; PRA; RFLP;  
 KW PCR-restriction fragment length polymorphism analysis; ds.

XX Mycobacterium bovis.

PN W0200131061-A1.

PD 03-MAY-2001.

XX 27-OCT-2000; 2000MO-KR01223.

PR 27-OCT-1999; 99KR-0046795.

PA (ERUM-) ERUME BIOTECH CO LTD.

PI Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;

DR WPI: 2001-300520/31.

XX New DNA fragments from the rpoB gene of mycobacteria, useful for  
 PT diagnosis and identification of many mycobacterial species by  
 PT restriction fragment length polymorphism -

PS Claim 1: Page 45; 50pp; English.

CC The present sequence for Mycobacterium bovis rpoB gene  
 CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05242) from  
 CC various Mycobacterial species. These rpoB gene fragments can be used  
 CC in the diagnosis and identification of Mycobacterium species using a  
 CC novel PCR-restriction fragment length polymorphism analysis (PRA)  
 CC method. The method comprises obtaining a restriction fragment length  
 CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,  
 CC amplifying and digesting the DNA fragment from the microorganism to  
 CC be identified and comparing the RFLP patterns from the known rpoB gene  
 CC fragments with the unidentified fragment. The rpoB gene fragments

CC are useful to identify a wide range of Mycobacterium species, e.g. for  
 CC diagnosis or to obtain epidemiological and pathogenesis information for  
 CC selection of appropriate therapies, including M. tuberculosis, M. leprae  
 CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected  
 CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene  
 CC fragments is rapid, precise, simple and cost effective (only 1 PCR  
 CC required), and can differentiate between many species in a single  
 CC experiment, including those difficult to distinguish by usual biochemical  
 CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for  
 CC detecting specific Mycobacterial species.

SO Sequence 208 BP; 46 A; 62 C; 67 G; 33 T; 0 other;

Query Match 80.0%; Score 166.4; DB 22; Length 208;  
 Best Local Similarity 87.5%; Pred. No. 1.9e-28;  
 Matches 182; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

OY 1 TCAGAGGAAGCGCTACGACCTGGCCCGCGGTGGCGCTACAGGTCAACAGAAAGCTGG 60  
 DB 1 TCAGAGGAAGCGCTACGACCTGGCCCGCGGTGGCGCTACAGGTCAACAGAAAGCTGG 60  
 OY 61 GCCTGAACCGCCGCGCATCAGCTGCTGCTACCTCTACCGAGAGAGAGCTGCGCA 120  
 DB 61 GCCTGATGTGCGCGCATCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120  
 OY 121 CCATCGAGTACCTGCTGCGCTGCGACGAGGCGGACAGCAGTACGCTGCGCGCGCG 180  
 DB 121 CCATCGAATATCTGCTGCGCTGCGACGAGGCGGACAGCAGTACGCTGCGCGCGCG 180  
 OY 181 TCGAGGTCCCGGTGAGGTGAGCAGCAT 208  
 DB 181 TCGAGGTCCCGGTGAGGTGAGCAGCAT 208

RESULT 12  
 AAT09676  
 ID AAT09676 standard; DNA: 970 BP.

AC AAT09676;

DT 15-OCT-1996 (first entry)

DE Mycobacterium tuberculosis rpoB gene DNA sequence.

XX Tuberculosis; disease diagnosis; oligonucleotide; DNA primer; PCR;  
 KW polymerase chain reaction; DNA amplification; rpoB locus; TB; ss.

XX Mycobacterium tuberculosis.

FT Key Location/Qualifiers

FT primer\_bind 10..27

FT /tag= a primer FENLFP\*

FT /note= b primer DDIDHL\*

FT /tag= c primer DDIDH\*

FT /tag= d primer DDIDH\*

FT /tag= e primer rpo95\*

FT /tag= f primer rpo105\*

FT /tag= g primer KY290\*

FT /tag= h M. tuberculosis signature nucleotide\*

FT misc-feature 433..434

FT /tag= h

FT misc-feature

FT

	Query Match	80.0%	Score 166.4;	DB 17;	Length 970;
	Best Local Similarity	87.5%;	Pred. No. 1,9e-28;		
	Matches 182; Conservative	0;	Mismatches 26;	Indels 0;	Gaps 0
OY	1 TCAAGAGAAAGCGGTACACACTGGGCCCCGGGGCCCTTACAAAGTAAACAAGACTG 60				
Db	26 TCAGAAGAGAGAGGCATACACCCTGGGCCCGCTTGCTTAAAGTCAACAACAACTCG 85				
OY	61 GCCTGAAGCGGGCCAGCCGATTCACGTGTGTAATTGACCGAGAAAGCTGTGCGCA 120				
Db	86 GGGTCATAGTCGGGAGGCCCATCACGTGTGACAGCTGTGACCGAAGAAGAGCTGTGCGCA 145				
OY	121 CCATGAGTACCTGGTGGCCTGTGACGAGGGCCAAGACACAGATGACCGTCCC GGCG 180				
Db	146 CCAATGATATATCTGGTCCGCTTGACAGAGGTACAGACCAGATGATGACCGTTCCGGGCGCG 205				
OY	181 TCGAGGTCCCGGTGAGAGTGCAGCAGAT 208				
Db	206 TCGAGGTCCCGGTGAAACCGACAGCAT 233				
<hr/>					
	RESULT 13				
	AAH51976				
ID	AAH51976 standard; DNA; 3519 BP.				
XX	AAH51976;				
AC					
XX					
DT	04-SEP-2001 (first entry)				
XX					
De	Mycobacterium tuberculosis potential drug target gene SEQ ID 30.				
KW	Drug target; growth; organism viability; characterisation; ds.				
XX	Mycobacterium tuberculosis.				
OS	MO2001J35317-AI.				
PN					
XX	17-MAY-2001.				
PD					
XX	13-NOV-2000; 2000WO-USJ1152.				
PF					
PR	12-NOV-1999; 99US-0165086.				
PR	12-NOV-1999; 99US-0165124.				
XX	01-FEB-2000; 2000UDS-0179531.				
PA	(REGC ) UNITV CALIFORNIA.				
XX					
PI	Eisenberg D, Rotstein SH, Marcotte EM;				
XX					
DR	WPI: 2001-329193/34.				
XX	P-PSDB: AAG81125.				
PT					
XX	Identifying nucleotide or polypeptide sequence for use as drug target,				
PT	involves providing algorithm that analyzes a functional relationship				
XX	between nucleotide or polypeptide sequences, and comparing the				
XX	sequences				
<hr/>					
	Disclosure: Page 68-69; 207pp: English.				
CC	This invention relates to a method for identifying a nucleotide or				
CC	polypeptide sequence that may be a drug target, or essential for growth				
CC	or viability of an organism. Polynucleotide sequences AAH51947 - AAH52092				
CC	represent DNA encoding proteins AAG81096 - AAG81241, Mycobacterium				
CC	tuberculosis proteins which are potential drug targets. The DNA and				
CC	protein sequences are used to illustrate the method of the invention. The				
CC	method involves providing an unknown nucleotide or polypeptide sequences,				
CC	and comparing it to a number of sequences along with at least one				
CC	algorithm capable of analysing a functional relationship between				
CC	nucleotide and polypeptide sequences. The method is useful for				
CC	characterising the function of nucleic acids and polypeptides that may be				
CC	useful as a target for a drug or essential for the growth or viability of				
CC	an organism.				

S0 Sequence 3519 BP; 675 A; 1081 C; 1183 G; 580 T; 0 other;  
Query Match 80.0%; Score 166.4; DB 22; Length 3519;  
Best Local Similarity 87.5%; Pred. No. 1.9e-28;  
Matches 182; Conservative 0; Mismatches 26; Indels 0; Gaps 0;  
OY 1 TCAAGGAGAGCGCTACGACCTGGCCGCGTGGCGCTACAAAGTCAACAGAGCTGG 60  
DB 884 TCAAGGAGAGCGCTACGACCTGGCCGCGTGGCGCTACAAAGTCAACAGAGCTGG 943  
OY 61 GCTCTGAACGCCGCCAGCGATCAGCTGCTGCACTCTGACCGAGGAGAGCTGTCGCCCA 120  
DB 944 GGTCTGATGTGCGGAGGCCATCAGCTGCTGACGCTGACCGAGAGAGCTGTCGGCCA 1003  
OY 121 CCATCGAGTACCTGTGTCGCCCTGCACGAGGCCACACCATGACCTGCCCGCGCGG 180  
DB 1004 CCATCGAATATCTGTCGCCCTGTCAGAGGTCACACCATGACCTGCCCGCGCGG 180  
OY 181 TCGAGTCCCGTTCGAGGTGGAGACAT 208  
DB 1064 TCGAGTCCCGTTCGAGGTGGAGACAT 1091  
RESULT 14  
ID AAH02079 standard; DNA; 3534 BP.  
XX AC AAH02079;  
XX  
XX 24-JUL-2001 (first entry)  
XX DE Mycobacterium tuberculosis nucleotide sequence SEQ ID NO:2072.  
XX KW Species specific; genus specific; family specific; probe; detection;  
KW identification; algal; archaeal; bacterial; fungal; parasitological;  
KW microorganism; diagnosis; translation elongation factor Tu; toxin;  
KW translation elongation factor G; RecA recombinase; resistance;  
KW catalytic subunit of proton-translocating ATPase; antimicrobial;  
XX vaccine; primer; ds.  
XX OS Mycobacterium tuberculosis.  
XX  
XX PN WO200123604-A2.  
XX PD 05-APR-2001.  
XX PF 28-SEP-2000; 2000WO-CA01150.  
XX PR 28-SEP-1999; 99CA-2283458.  
XX PS 19-MAY-2000; 2000CA-2307010.  
XX PA (INFE-) INFECTIO DIAGNOSTIC (IDI) INC.  
XX PI Bergeron MG, Boissinot M, Huletsky A, Menard C, Ouellette M;  
XX Picard FJ, Roy PH;  
XX WPI; 2001-245006/25.  
XX  
XX Nucleic acid sequences are used to generate universal probes and  
XX primers which can be used to identify and detect the presence of algal,  
XX archaeal, bacterial, fungal and parasitological species in a test sample -  
XX  
XX Disclosure; Page 1478-1479; 1580pp; English.  
XX  
XX The present invention describes a method for generating a repertoire of  
XX nucleic acids of tuf, fus, atpD and/or recA genes from which probes  
XX and/or primers are derived. The method comprises amplifying the nucleic  
XX acids of determined algal, archaeal, bacterial, fungal and parasitological  
XX species with a combination of defined primer pairs. The method can be  
XX used for producing probes and/or primers for detecting one or more  
XX related microorganisms e.g. algae, archaea, bacteria, fungi and  
XX parasites, for universal detection and for specific and ubiquitous  
XX detection and identification of an algal, archaeal, bacterial, fungal

CC and parasitological species, genus, family and group. A nucleic acid (1)  
CC obtained using the method of the invention can be used for the universal  
CC detection of any bacterium, fungus or parasite in a sample and for the  
CC detection of at least one antimicrobial agent resistance gene or at  
CC least one toxin gene. hexA nucleic acids are used for the specific and  
CC ubiquitous detection and for identification of Streptococcus pneumoniae.  
CC (1) can be used to design a therapeutic agent which is effective against  
CC microorganisms. Microbial species or genus or family or phylum or group  
CC which can be detected include Abiotrophia adiacens, Bordetella sp.,  
CC Corynebacterium sp., Enterobacteriaceae group, Escherichia coli,  
CC Mycobacteriaceae family, Pseudomonas group, Streptococcus sp.,  
CC Neisseria gonorrhoeae and Staphylococcus sp.. Using DNA based tests  
CC provides faster results than substrate specificity tests as results can  
CC be determined in an hour and improved accuracy is also achieved.  
CC AAH00010 to AAH002304 represent nucleotide sequences and primers/probes  
XX which are given in the exemplification of the present invention.  
SQ Sequence 3534 BP; 679 A; 1081 C; 1188 G; 586 T; 0 other;  
Query Match 80.0%; Score 166.4; DB 22; Length 3534;  
Best Local Similarity 87.5%; Pred. No. 1.9e-28;  
Matches 182; Conservative 0; Mismatches 26; Indels 0; Gaps 0;  
OY 1 TCAAGGAGAGCGCTACGACCTGGCCGCGTGGCGCTACAAAGTCAACAGAGCTGG 60  
DB 902 TCAAGGAGAGCGCTACGACCTGGCCGCGTGGCGCTACAAAGTCAACAGAGCTGG 961  
OY 61 GCTCTGAACGCCGCCAGCGATCAGCTGCTGCACTCTGACCGAGGAGAGCTGTCGCCCA 120  
DB 962 GGTCTGATGTGCGGAGGCCATCAGCTGCTGACGCTGACCGAGAGAGCTGTCGCCCA 1021  
OY 121 CCATCGAGTACCTGTGTCGCCCTGCACGAGGCCACACCATGACCTGCCCGCGCGG 180  
DB 1022 CCATCGAATATCTGTCGCCCTGTCAGAGGTCACACCATGACCTGCCCGCGCGG 1081  
OY 181 TCGAGTCCCGTTCGAGGTGGAGACAT 208  
DB 1082 TCGAGTCCCGTTCGAGGTGGAGACAT 1109  
RESULT 15  
AAH74651  
ID AAA74651 standard; DNA; 3853 BP.  
XX AC AAA74651;  
XX  
XX 06-DEC-2000 (first entry)  
XX DE Mycobacterium tuberculosis rpoB gene.  
XX KW Mycobacterium tuberculosis; rpoB; RNA polymerase beta subunit;  
XX rifampin resistance; mutation detection; ds.  
XX OS Mycobacterium tuberculosis.  
XX PN WO200043546-A2.  
XX PD 27-JUL-2000.  
XX PF 20-DEC-1999; 99WO-US30377.  
XX PR 19-JAN-1999; 99US-0233996.  
XX PA (DADE-) DADE BEHRING INC.  
XX PI Liu YP, Kurn N;  
XX WPI; 2000-524243/47.  
XX Method for detecting drug resistance in a strain of an organism,  
XX particularly for detecting rifampin resistance in Mycobacterium  
XX tuberculosis -

**Example 1; Fig 4; 86pp; English.**

The present sequence is the rpoB gene from Mycobacterium tuberculosis.  
Rifampin resistance is largely associated with point mutations  
localised in a small core region of 81 base pairs in the rpoB gene, which  
encodes the RNA polymerase beta subunit. To detect a mutation, a complex  
is formed comprising a first sequence representing the predetermined  
region of the gene of the organism and a second sequence representing the  
corresponding region of the gene of the wild type organism in double  
stranded form. Each member of at least one pair of non-complementary  
strands within the complex has a label. The association of the labels in  
the complex is related to the presence of the mutation. The presence of  
the mutation is related to the drug resistance of the strain.

Query Match 80.0%; Score 166.4; DB 21; Length 3853;  
Best Local Similarity 87.5%; Pred. No. 1.9e-28;  
Matches 182; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY	1	TC	AAGGAGAGCGCTACGACCTGGCCCGCTACAAAGGTCAACAAGAGCTGG	60
Db	1477	TC	AAGGAGAGCGCTACGACCTGGCCCGCTACAAAGGTCAACAAGAGCTGG	1536
QY	61	GC	CTGAACCGCCGACGCGGATCAGCTCGTCGACTCTGACCGGAGGAAGAGCTGTCGC	120
Db	1537	GG	CTGATGTGGCGAGGCCATCAGCTGTGTCGAGCTGACCGAAGAAGACGCTGTGGCCA	1596
QY	121	CC	ATCAGGTACCTGGTGGCGCCTGCACGAGGCCACGACGATGACCGTCCCGCGGCGC	180
Db	1597	CC	ATCGAATATCTGCTCGCTTGACAGAGGTCAGACCACCATGACCGTTCCGGGGCGG	1656
QY	181	TC	CAGTCCCGCTCGAGTGGACGACAT	208
Db	1657	TC	CAGTCCCGTGGAAACCGACGACAT	1684

Search completed: November 12, 2002, 16:50:23  
Job time : 143.222 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 12, 2002, 16:50:46 ; Search time 636.023 Seconds  
(without alignments)  
9517.553 Million cell updates/sec

Title: US-09-697-123B-19  
Perfect score: 208  
Sequence: 1 tcaagagagaagcgctacgac.....ccggctcgagggtggaagacat 208

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues  
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

- 1: gb\_ba.\*
- 2: gb\_htg.\*
- 3: gb\_in.\*
- 4: gb\_om.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sts.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vi.\*
- 15: em\_ba.\*
- 16: em\_fun.\*
- 17: em\_hum.\*
- 18: em\_in.\*
- 19: em\_mu.\*
- 20: em\_om.\*
- 21: em\_or.\*
- 22: em\_ov.\*
- 23: em\_pat.\*
- 24: em\_ph.\*
- 25: em\_pl.\*
- 26: em\_ro.\*
- 27: em\_sts.\*
- 28: em\_un.\*
- 29: em\_vi.\*
- 30: em\_htg\_hum.\*
- 31: em\_htg\_inv.\*
- 32: em\_htg\_other.\*
- 33: em\_htg\_mus.\*
- 34: em\_htg\_pln.\*
- 35: em\_htg\_rod.\*
- 36: em\_htg\_mam.\*
- 37: em\_htg\_vrt.\*
- 38: em\_sy.\*
- 39: em\_htgo\_hum.\*
- 40: em\_htgo\_mus.\*
- 41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	171.2	82.3	3752	1	MSU24494	U24494 Mycobacteri
2	166.4	80.0	610	1	MTU318818	AJ318818 Mycobacte
3	166.4	80.0	610	1	MTU318819	AJ318819 Mycobacte
4	166.4	80.0	616	1	MTU318813	AJ318813 Mycobacte
5	166.4	80.0	618	1	MTU318815	AJ318815 Mycobacte
6	166.4	80.0	618	1	MTU318817	AJ318817 Mycobacte
7	166.4	80.0	633	1	MTU318814	AJ318814 Mycobacte
8	166.4	80.0	637	1	MTU318816	AJ318816 Mycobacte
9	166.4	80.0	639	1	MTU318821	AJ318821 Mycobacte
10	166.4	80.0	970	6	I50706	I50706 Sequence 1
11	166.4	80.0	3534	6	AX111339	AX111339 Sequence
12	166.4	80.0	3853	1	MTU12205	U12205 Mycobacteri
13	166.4	80.0	5084	1	MSGRPOB	L37989 Mycobacteri
14	166.4	80.0	19352	1	AE006964	AE006964 Mycobacte
15	166.4	80.0	19770	1	WTC1376	Z35972 Mycobacteri
16	142.4	68.5	3447	6	AR067447	AR067447 Sequence
17	142.4	68.5	37617	1	MLB1790G	Z14314 M.leprae ge
18	142.4	68.5	348950	1	MLEPRTN7	AL583923 Mycobacte
19	109.2	52.5	3941	1	AF242549	AF242549 Amycolato
20	88	42.3	3495	6	AX120631	AX120631 Sequence
21	88	42.3	328050	1	AP005275	AP005275 Corynebac
22	88	42.3	349980	6	AX127144	AX127144 Sequence
23	86.8	41.7	32923	1	SCDB2	AL160431 Streptomy
24	51.2	24.6	204050	1	AL646073	AL646073 Ralstonia
25	48.2	23.2	204050	1	AL646070	AL646070 Ralstonia
26	47.4	22.8	32748	1	AB070951	AB070951 Streptomy
27	47.2	22.7	7954	1	TAQ19223	V19223 Thermus aqu
28	46.6	22.4	1416	1	SRRECAGEN	X94233 S.rimosus r
29	46.2	22.2	20054	1	SCCB12	AL391588 Streptomy
30	45.8	22.0	329100	1	SME591787	AL591787 Sinorhizo
31	45.6	21.9	1271	1	SVU04837	U04837 Streptomyce
32	45	21.6	134816	2	CNS08C9R	AL772413 Oryza sat
33	45	21.6	135378	2	CNS07YP9	AL713941 Oryza sat
34	44.8	21.5	1952	1	SCU42212	U42212 Streptomyce
35	44.4	21.3	1782	14	AF171937	AF171937 Suid herp
36	44.2	21.3	1734	6	I04901	I04901 Sequence 6
37	44.2	21.3	12112	1	AE005722	AE005722 Caulobact
38	44.2	21.3	92509	1	AL646086	AL646086 Ralstonia
39	44.2	21.3	346294	1	AP002999	AP002999 Mesorhizo
40	44	21.2	1490	1	SARECAG	Z30324 S.ambofacie
41	44	21.2	3488	1	SCWHIB	X62287 S.coelicolo
42	44	21.2	37898	1	SCE34	AL353862 Streptomy
43	43.8	21.1	1440	8	AY034826	AY034826 Curvulari
44	43.8	21.1	2206	3	AY102669	AY102669 Drosophila
45	43.8	21.1	9302	1	SGR243674	AJ243674 Streptomy

ALIGNMENTS

RESULT 1  
MSU24494  
LOCUS MSU24494 3752 bp DNA linear BCT 02-MAR-2000  
DEFINITION Mycobacterium smegmatis DNA polymerase (rpoB) gene, complete cds.  
ACCESSION U24494  
VERSION U24494.1 GI:790347  
KEYWORDS  
SOURCE  
ORGANISM Mycobacterium smegmatis.  
Mycobacterium smegmatis  
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
Actinomycetales; Corynebacterineae; Mycobacteriaceae;  
Mycobacterium.  
REFERENCE 1 (bases 1 to 3752)  
AUTHORS Hetherington,S.V., Watson,A.S. and Patrick,C.C.  
TITLE Sequence and analysis of the rpoB gene of Mycobacterium smegmatis

JOURNAL	Antimicrob. Agents Chemother. 39 (9), 2164-2166 (1995)
MEDLINE	96050766
PUBMED	8340740
REFERENCE	2 (bases 1 to 3752)
AUTHORS	Hetherington,S.V.
TITLE	Direct Submission
JOURNAL	Submitted (11-APR-1995) Seth V. Hetherington, Infectious Diseases, St. Jude Children's Research Hospital, 332 N. Lauderdale, Memphis, TN 38101, USA
FEATURES	Location/Qualifiers source 1..3752 /organism="Mycobacterium smegmatis" /db_xref="taxon:1772" 194..3703 /gene="rpoB" 194..3703 /gene="rpoB" /codon_start=1 /evidence-experimental /transl_table=11 /product="DNA polymerase" /protein_id="AA091426.2" /db_xref="GI:7144498" /translation="MLEGICLAVSSQSKNAITNSVRGPNVSGFAKLREPLEVRGL LDVQTSFELWYDGRVROQAIDRGEPNVGVLEAEISPIEDFGSMSSLSGSDPR FEVKASVDKMDTYAALFTVAEPIINNTGEIKSOTVMGDFPMWTKEGTFFIN GTERVVSVQLDETDIDKTEKTLKSVKIPGRGAAGFEEDDKRDYVGVRID RKRRPVTLLKALGMTNEOIVREFGFSEIMMGTTLEKDTTSCGDEALLDIYRLRGE PTPKESAOTLLLENLFKKRYDLARVGYKNKLGILNAGKPITSSTLTBEDVVATIE YLVLHEGOSTMTVPGGVEFAVESHDIDHFNRLRTVGEIQIRVLRSMERVVR ERTMQDVAITQTILNIRPVAAIKEFTQSLSOFMDQNPLSLGTHKRRLSALG PGLSRERAGLEVDRHPSHRGMCPIETPEGNIGLSLVYRVNPFGEIETPYR KVQEVGTDIDYLADEEDHRVVAQANSPTDENRFTEDRVVVRKKGVEFVSADQ VDMYDSRQMSVATAMIPFLHDDANRALGMORAOVPLVRSEAPLVGTGMELR AADVWTSKHTGVIIEVSADYITVMADGTROS YRLRFKPARNSHCTCANQPIDA GORVEAGOVADPGCTONGENALKNLLVAMPHEGNYVEDAIIISNRLVEEDVLISI HIEEHIDARDPKLCAEBITVPIKLSDEVIALDDEFRIVRAEVRDGELLVGKVTTP KGETELTEERLLAIRAFEGAREVRDTSKLPHGSESKVIGIRVFREDDEDLPVGN ELVRVYPAKRKISGDKLARKHNGKVGIGILPVEDMPFLPDGTPVDIILNTHGCRV VELAGLLSTLPNRDGEYVNADGKATLFDGRSSGEPFPYPVTYGVYILKLHLVDD KIARSTGPSYMITOOLPGKAQFGQRFGMECWAMOAYGAAYTLOELLTIKSDDTV GRVKYEATVGENIPEGPISPESFKVLLKELQSLCLNVEVLSDDGRAENRDGDDDL ERAAANLGINLSRNESASVEDLA"
BASE COUNT	706 a 1225 c 1210 g 611 t
ORIGIN	
Query Match	82.3%; Score 171.2; DB 1; Length 3752;
Best Local Similarity	88.9%; Pred. No. 4.6e-19;
Matches 185; Conservative	0; Mismatches 23; Indels 0; Gaps 0;
QY	1 TCAAGGAGAAGCGCTACGACCTGGCCCGCTGGCCGCTACAAGTCACAAAGAGCTGG 60
Db	1068 TCAAGGAGAAGCGCTACGACCTGGCCCGTGTCGGCGCTTACCAAGGTCAACAAAGAGCTGG 1127
QY	61 GCCTGAACGCCGCCACCGCATCAGCTCGTGACTCTTGACCAGGAGGACACGCTCGTCGC 120
Db	1128 GCCTGAACGCCGCCAACGCCGATCACACGCTCGACCTGACCCAGAGAGGACGCTCGTCG 1187
QY	121 CCATCGAGTACCTGGTCCCTTCACAGGAGGCGCACACCATGACCGTCCCCCGGGCGG 180
Db	1188 CCATCGAGTACCTGGTGGCTCTGCACAGGGGTGACACCTCGATGACCGTCCC CGGTGCG 1247
QY	181 TCGAGGTCCCGGTCGAGGTGGACGACAT 208
Db	1248 TCGAGTTCGCGTCGAGTCGCACGACAT 1275
RESULT 2	
MTU318818	
LOCUS	610 bp DNA circular BCT 09-AUG-2002
DEFINITION	Mycobacterium tuberculosis partial rpoB gene for RNA polymerase beta subunit, isolate 1415-97.
ACCESSION	MTU318819
VERSION	1
KEYWORDS	RNA polymerase beta subunit; rpoB gene.
SOURCE	Mycobacterium tuberculosis.
ORGANISM	Bacteria; Actinobacteria; Actinomycetales:

Corynebacterineae; Mycobacteriaceae; Mycobacterium tuberculosis complex.

REFERENCE  
AUTHORS  
TITLE  
1 Herrera, L., Jimenez, M.S. and Saez, J.A.  
Molecular analysis of rifampin-resistant *Mycobacterium tuberculosis*  
isolated in Spain (1996-2001). Description of new alleles into *rpoB*  
gene and review

JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
Submitted (07-AUG-2001) Herrera L., Servicio Bacteriología, Centro  
Nacional Microbiología., Ctra. Majadahonda-Pozuelo, Km 2.5,  
Majadahonda, Madrid, 28220. SPAIN

```

FEATURES
source
Location/Qualifiers
1. .610
/organism="Mycobacterium tuberculosis"
/isolate="1417-97"
/db xref="taxon.1773"

```

```

gene
1. .610
/ob_xref= taxon:1773
/gene="rpoB"
CDS
<1..>610
/ncore="rpoB"

```

```
/gene="rpoB"  
/codon_start=1  
/transl_table=11  
/product="RNA polymerase beta subunit"  
/protein_id="CAC87036.1"  
/def="rpoB; CAC87036.1"
```

```

/ob_xref="GI:22208415"
/translation="LDIYKRLRPPETKESAQTLLLENLFFKEKRYDLARVGRKYKN
KLGKHGPEPTSTLTEDDVATIEYLRHHEQQTMTVPGGEVPEVETDDIDHFCNR
RUTVCELGTONQIRGMSRMRVERMTQDVEAITPOTLINIRPVAAIKEFFCTS
QPSQFGNQNPISLTHKRLRLSALGPGGLSRERAGLEVRDVHP"
122 a 191 c 202 g 95 t
BASE COUNT
ORIGIN

```

QY	181	TCGAGGTC	CCGGT	TCGAGG	TCGAGC	GACAT	208
Db	260	TCGAGGTC <th>CCGGT</th> <td>GGAAAC</td> <td>CGAC <td>GACAT <td>287</td> </td></td>	CCGGT	GGAAAC	CGAC <td>GACAT <td>287</td> </td>	GACAT <td>287</td>	287

RESULT 4	LOCUS	DEFINITION
MTU318813	616 bp	DNA
	Mycobacterium tuberculosis partial	gene for RNA polymerase
	beta subunit, isolate 1763-97.	

ACCESSION	AJ318813
VERSION	AJ318813.1
KEYWORDS	RNA polymerase beta subunit; rpoB gene.
SOURCE	Mycobacterium tuberculosis.
ORGANISM	Mycobacterium tuberculosis. Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium tuberculosis complex.

REFERENCE	AUTHORS	TITLE
1	Herrera, L., Jimenez M.S. and Saez, J.A.	Molecular analysis of rifampin-resistant <i>Mycobacterium tuberculosis</i> isolated in Spain (1996-2001). Description of new alleles into <i>rpoB</i>

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

**JOURNAL  
REFERENCE  
AUTHORS**

FEATURES SOURCE

gene

CDS

**LOCUS**  
**DEFINITION**

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

REFERENCE	AUTHORS	TITLE
1	...	...
2	...	...
3	...	...
4	...	...
5	...	...
6	...	...
7	...	...
8	...	...
9	...	...
10	...	...
11	...	...
12	...	...
13	...	...
14	...	...
15	...	...
16	...	...
17	...	...
18	...	...
19	...	...
20	...	...
21	...	...
22	...	...
23	...	...
24	...	...
25	...	...
26	...	...
27	...	...
28	...	...
29	...	...
30	...	...
31	...	...
32	...	...
33	...	...
34	...	...
35	...	...
36	...	...
37	...	...
38	...	...
39	...	...
40	...	...
41	...	...
42	...	...
43	...	...
44	...	...
45	...	...
46	...	...
47	...	...
48	...	...
49	...	...
50	...	...
51	...	...
52	...	...
53	...	...
54	...	...
55	...	...
56	...	...
57	...	...
58	...	...
59	...	...
60	...	...
61	...	...
62	...	...
63	...	...
64	...	...
65	...	...
66	...	...
67	...	...
68	...	...
69	...	...
70	...	...
71	...	...
72	...	...
73	...	...
74	...	...
75	...	...
76	...	...
77	...	...
78	...	...
79	...	...
80	...	...
81	...	...
82	...	...
83	...	...
84	...	...
85	...	...
86	...	...
87	...	...
88	...	...
89	...	...
90	...	...
91	...	...
92	...	...
93	...	...
94	...	...
95	...	...
96	...	...
97	...	...
98	...	...
99	...	...
100	...	...

JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL

```
Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5,
Majadahonda. Madrid. 28220, SPAIN
Location/Qualifiers
1..618
/organism="Mycobacterium tuberculosis"
/isolate="2540-97"
/db_xref="taxon:1773"
1..618
/gene="rpoB"
<1..>618
/codon_start=1
/transl_table=11
/product="RNA polymerase beta subunit"
/protein_id="CAC87032.1"
/db_xref="GI:22208407"
/translation="TDEALLDIYKLRPGEPPTKESAOITLLENLFFKPKRYDLARVGR
YKNNKLGHLHVGEPITSSLTLEEDVYATIEYLRLHGGQTTMTVPGGVEVPVETDDID
HFGNRLRTYVGLIQIRVGNRMERVRMTTQDVEAITPTLIRPVAIAKE
FFGTQSLSQFMQNNPLSLGTHKRRLALPGGLSRERAGLEVRDV"
BASE COUNT 124 a 190 c 207 g 97 t
ORIGIN
Query Match 80.0%; Score 166.4; DB 1; Length 618;
Best Local Similarity 87.5%; Pred. No. 3.8e-18;
Matches 182; Conservative 0; Mismatches 26; Indels 0; Gaps 0;
QY 1 TCAAGGAGAACGCTACGACCTGGCCGCGTGGCGCTACAGGTCAACAAGAGCTGG 60
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 95 TCAAGGAGAACGCTACGACCTGGCCGCGTGGCGCTACAGGTCAACAAGAGCTGG 154
QY 61 GCCTGAACGCCGCCAGCCGATCAGCTGCTGCGTCTGACCGAGGAGAGCTGCTGCCA 120
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 155 GCCTGATGTCGGCGAGCCATCACGCTGCTGCGAGCTGACCGAAGAGAGCTGCTGCCA 214
QY 121 CCATCGAGTACCTGTCGCTGCACGAGGCCAGACGATGACCGTCCCGCGCGCG 180
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 215 CCATCGAATATCTGTCGCTGTCACGAGGTCAGACGATGACCGTCCCGCGCGCG 274
QY 181 TCGAGTCCCGTTCGAGGTGGAGACAT 208
||||| ||||| || |||||
Db 275 TCGAGTCCCGTGGAAACCGAGACAT 302

RESULT 6
MTU318817
LOCUS
DEFINITION
Mycobacterium tuberculosis partial rpoB gene for RNA polymerase
beta subunit, isolate 2348-98.
ACCESSION
AJ318817
VERSION
GI:22208410
KEYWORDS
RNA polymerase beta subunit; rpoB gene.
SOURCE
Mycobacterium tuberculosis.
ORGANISM
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
tuberculosis complex.
1
REFERENCE
Herrera, L., Jimenez, M.S. and Saez, J.A.
Molecular analysis of rifampin-resistant Mycobacterium tuberculosis
isolated in Spain (1996-2001). Description of new alleles into rpoB
gene and review
Unpublished
2 (bases 1 to 618)
Herrera, L.
Direct Submission
Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro
Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5,
Majadahonda. Madrid. 28220, SPAIN
Location/Qualifiers
1..618
/organism="Mycobacterium tuberculosis"
/isolate="2348-98"
FEATURES
source
1..618
/gene="rpoB"
<1..>618
/codon_start=1
/transl_table=11
/product="RNA polymerase beta subunit"
/protein_id="CAC87032.1"
/db_xref="GI:22208407"
/translation="TDEALLDIYKLRPGEPPTKESAOITLLENLFFKPKRYDLARVGR
YKNNKLGHLHVGEPITSSLTLEEDVYATIEYLRLHGGQTTMTVPGGVEVPVETDDID
HFGNRLRTYVGLIQIRVGNRMERVRMTTQDVEAITPTLIRPVAIAKE
FFGTQSLSQFMQNNPLSLGTHKRRLALPGGLSRERAGLEVRDV"
BASE COUNT 124 a 190 c 207 g 97 t
ORIGIN
Query Match 80.0%; Score 166.4; DB 1; Length 618;
Best Local Similarity 87.5%; Pred. No. 3.8e-18;
Matches 182; Conservative 0; Mismatches 26; Indels 0; Gaps 0;
QY 1 TCAAGGAGAACGCTACGACCTGGCCGCGTGGCGCTACAGGTCAACAAGAGCTGG 60
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 95 TCAAGGAGAACGCTACGACCTGGCCGCGTGGCGCTACAGGTCAACAAGAGCTGG 154
QY 61 GCCTGAACGCCGCCAGCCGATCAGCTGCTGCGTCTGACCGAGGAGAGCTGCTGCCA 120
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 155 GCCTGATGTCGGCGAGCCATCACGCTGCTGCGAGCTGACCGAAGAGAGCTGCTGCCA 214
QY 121 CCATCGAGTACCTGTCGCTGCACGAGGCCAGACGATGACCGTCCCGCGCGCG 180
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 215 CCATCGAATATCTGTCGCTGTCACGAGGTCAGACGATGACCGTCCCGCGCGCG 274
QY 181 TCGAGTCCCGTTCGAGGTGGAGACAT 208
||||| ||||| || |||||
Db 275 TCGAGTCCCGTGGAAACCGAGACAT 302

RESULT 7
MTU318814
LOCUS
DEFINITION
Mycobacterium tuberculosis partial rpoB gene for RNA polymerase
beta subunit, isolate 1058-97.
ACCESSION
AJ318814
VERSION
GI:22208404
KEYWORDS
RNA polymerase beta subunit; rpoB gene.
SOURCE
Mycobacterium tuberculosis.
ORGANISM
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
tuberculosis complex.
1
REFERENCE
Herrera, L., Jimenez, M.S. and Saez, J.A.
Molecular analysis of rifampin-resistant Mycobacterium tuberculosis
isolated in Spain (1996-2001). Description of new alleles into rpoB
gene and review
Unpublished
2 (bases 1 to 633)
Herrera, L.
Direct Submission
Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro
Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5,
Majadahonda. Madrid. 28220, SPAIN
Location/Qualifiers
1..633
/organism="Mycobacterium tuberculosis"
/isolate="1058-97"
/db_xref="taxon:1773"
1..633
/gene="rpoB"
<1..>633
/codon_start=1
FEATURES
source
1..633
/organism="Mycobacterium tuberculosis"
/isolate="1058-97"
/db_xref="taxon:1773"
1..633
/gene="rpoB"
<1..>633
/codon_start=1
```

IDHFGNRLRTVCGELIQNIQVGRSMRERVVREBMTQDVEAITPOTLINIRPVAAI					
KEPFGTSOLFQNDONNPLSLGPHKRRMPALGPGGLSRAGLEVDRVHPSH*					
BASE COUNT	128 a	198 c	210 g	101 t	
ORIGIN					
Query Match	80.0%	Score 166.4;	DB 1;	Length 637;	
Best Local Similarity	87.5%;	pred. No. 3.7e-18;			
Matches 182;	Conservative 0;	Mismatches 26;	Indels 0;	Gaps 0;	
Qy 1	TCAAGGACAAGCGGTACGACCTGGCCCGCTGGCCGCTACAGGTCACACGAAGCTGG 60				
Db 101	TCAAGGAGAAGCGGTACGACCTGGCCCGCTGGCGTATGAGGTCAACAAGAAGCTCG 160				
Qy 61	GCTTGAACCGCGCCAGCGGATCAGCTGCTGCAGTCTTGACGAGGAAGACGTCCTCGCCA 120				
Db 161	GGCTGCATGTTCGGCGAGGCCATCAGCTGCTGCAGCGCTGACCGGAAGAAGACGTCGTGGCCA 220				
Qy 121	CCATCGAGTACTGGTGGCGCTTGCACGAGGCCACAGCATGACGTCCTCCCGGGCGCG 180				
Db 221	CCATCGAATATCTGGTTCGGCTTGCACGAGGTTCAGACACGATACCGCTTCCGGGGCGCG 280				
Qy 181	TCGAGGTCCCGGTTCGAGGTGGACGACAT 208				
Db 281	TCGAGGTCCCGGTGGAAACCGACGACAT 308				
RESULT 9					
MTU318821					
LOCUS	639 bp DNA circular BCT 09-AUG-2002				
DEFINITION	Mycobacterium tuberculosis partial rpoB gene for RNA polymerase beta subunit, isolate 1071-98.				

ACCESSION	AJ3118821	
VERSION	AJ3118821.1	GI:22208418
KEYWORDS	RNA polymerase beta subunit; rpoB gene.	
SOURCE	Mycobacterium tuberculosis.	
ORGANISM	Mycobacterium tuberculosis	
	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium tuberculosis complex.	
REFERENCE	1	
AUTHORS	Herrera,L., Jimenez,M.S. and Saez,J.A.	
TITLE	Molecular analysis of rifampin-resistant Mycobacterium tuberculosis isolated in Spain (1996-2001). Description of new alleles into rpoB gene and review	
JOURNAL	Unpublished	
REFERENCE	2 (bases 1 to 639)	
AUTHORS	Herrera,L.	
TITLE	Direct Submission	
JOURNAL	Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5, Majadahonda. Madrid. 28220, SPAIN	
FEATURES	Location/Qualifiers	
source	1..639	
	/organism="Mycobacterium tuberculosis"	
	/isolate="10711-98"	
	/db_xref="taxon:1773"	
gene	1..639	
	/genes="rpoB"	
CDS	<1..>639	
	/genes="rpoB"	
	/codon_start=1	
	/transl_table=11	
	/product="RNA polymerase beta subunit"	
	/protein_id="CAC87038.1"	
	/db_xref="GI:22208419"	
	/translation="RRTDQALDIYKLRPGPPPTKESQTLLENLFFKFKKRYDLARYGRYKGNKLGHVGPITSTTEEDVAVTIEYLVRLHEGQTTMTVPGGVEVPVETDQIDHFGNRRLLRTVGLGELIQIOIRVGMSEMRVVRMTQDVEAITPTQLINIRPAVAALKEFFGTSQLSQFMVQNNPLSGLTOKRLLSALGPGGLSRERAGLEVDRVHPSHS"	
BASE COUNT	126 a	202 c 212 g 99 t
ORIGIN		
Query Match	80.0%;	Score 166.4; DB 1; Length 639;

Best Local Similarity 87.5%; Pred. No. 3.7e-18;  
Matches 182; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 1 TCAAGGAGAACGCTACGACCTGGCCCGCGTGGCGCGCTACAAGGTCAACAAGAGCTGG 60  
|||||  
Db 101 TCAAGGAGAACGCTACGACCTGGCCCGCGTGGCGCGCTATAAGGTCAACAAGAGCTCG 160  
QY 61 GCCTGAACCGCCGCGCAGCGATCAGCTGCTGACCTCTCACCAGGAGGACGCTGCTGCGCA 120  
|||||  
Db 161 GGCTGCATGTCGCGGAGCCCATCAGCTGCTGACGCTGACCGAAGACGCTGCTGGCCA 220  
QY 121 CCATCGAGTACCTGTGGCGCTGACGAGGCGGACGACGATGACCGTCCCGCGCGCG 180  
|||||  
Db 221 CCATCGAATATCTGCTCGCTTGCAGGAGGTTCAGACCACGATGACCGTTCGCGGCGCG 280  
QY 181 TCGAGGTCCCGTTCGAGGTGGACGACAT 208  
|||||  
Db 281 TCGAGGTCCCGTGGAAACCGACGACAT 308

RESULT 10  
150706  
LOCUS 150706 970 bp DNA linear PAT 07-OCT-1997  
DEFINITION Sequence 1 from patent US 5643723.  
ACCESSION 150706  
VERSION 150706.1 GI:2472409  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.

REFERENCE 1 (bases 1 to 970)  
AUTHORS Persing, D.H., Hunt, J.J., Young, K.Y., Feimlee, T.A., Roberts, G.D.  
and Whelan, A. Christian.  
TITLE Detection of a genetic locus encoding resistance to rifampin in  
mycobacterial cultures and in clinical specimens  
JOURNAL Patent: US 5643723-A 1 01-JUL-1997;  
FEATURES  
source Location/Qualifiers  
1. .970  
/organism="unknown"

BASE COUNT 182 a 302 c 330 g 156 t  
ORIGIN

Query Match 80.0%; Score 166.4; DB 6; Length 970;  
Best Local Similarity 87.5%; Pred. No. 3.5e-18;  
Matches 182; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 1 TCAAGGAGAACGCTACGACCTGGCCCGCGTGGCGCGCTACAAGGTCAACAAGAGCTGG 60  
|||||  
Db 26 TCAAGGAGAACGCTACGACCTGGCCCGCGTGGCGCGCTATAAGGTCAACAAGAGCTCG 85  
QY 61 GCCTGAACCGCCGCGCAGCGATCAGCTGCTGACCTCTGACCGAGGAGAGCTGCTGCGCA 120  
|||||  
Db 86 GGCTGCATGTCGCGGAGCCCATCAGCTGCTGACGCTGACCGAAGAGAGCTGCTGGCCA 145  
QY 121 CCATCGAGTACCTGTGGCGCTGACGAGGCGGACGACGATGACCGTCCCGCGCGCG 180  
|||||  
Db 146 CCATCGAATATCTGCTCGCTTGCAGGAGGTTCAGACCACGATGACCGTTCGCGGCGCG 205  
QY 181 TCGAGGTCCCGTTCGAGGTGGACGACAT 208  
|||||  
Db 206 TCGAGGTCCCGTGGAAACCGACGACAT 233

RESULT 11  
AX111339  
LOCUS AX111339 3534 bp DNA linear PAT 30-APR-2001  
DEFINITION Sequence 2072 from Patent WO0123604.  
ACCESSION AX111339  
VERSION AX111339.1 GI:13927631  
KEYWORDS  
SOURCE Mycobacterium tuberculosis.  
ORGANISM Mycobacterium tuberculosis.  
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;

Actinomycetales; Corynebacterineae; Mycobacteriaceae;  
Mycobacterium; Mycobacterium tuberculosis complex.  
1 (bases 1 to 3534)  
Bergeron, M.G., Boissinot, M., Huletsky, A., m Nard, C., Ouellette, M.,  
Picard, F.J. and Roy, P.H.  
Highly conserved genes and their use to generate probes and primers  
for detection of microorganisms  
Patent: WO 0123604-A 2072 05-APR-2001;  
Infectio Diagnostic (I.D.I.) INC. (CA)  
Location/Qualifiers  
1. .3534  
/organism="Mycobacterium tuberculosis"  
/strain="Rv"  
/db\_xref="taxon:1773"

BASE COUNT 679 a 1081 c 1188 g 586 t  
ORIGIN

Query Match 80.0%; Score 166.4; DB 6; Length 3534;  
Best Local Similarity 87.5%; Pred. No. 2.8e-18;  
Matches 182; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 1 TCAAGGAGAACGCTACGACCTGGCCCGCGTGGCGCGCTACAAGGTCAACAAGAGCTGG 60  
|||||  
Db 902 TCAAGGAGAACGCTACGACCTGGCCCGCGTGGCGCGCTATAAGGTCAACAAGAGCTCG 961  
QY 61 GCCTGAACCGCCGCGCAGCGATCAGCTGCTGACCTCTGACCGAGGAGAGCTGCTGCGCA 120  
|||||  
Db 962 GGCTGCATGTCGCGGAGCCCATCAGCTGCTGACGCTGACCGAAGAGAGCTGCTGGCCA 1021  
QY 121 CCATCGAGTACCTGTGGCGCTGACGAGGCGGACGACGATGACCGTCCCGCGCGCG 180  
|||||  
Db 1022 CCATCGAATATCTGCTCGCTTGCAGGAGGTTCAGACCACGATGACCGTTCGCGGCGCG 1081  
QY 181 TCGAGGTCCCGTTCGAGGTGGACGACAT 208  
|||||  
Db 1082 TCGAGGTCCCGTGGAAACCGACGACAT 1109

RESULT 12  
MTU12205  
LOCUS MTU12205 3853 bp DNA linear BCT 02-MAR-2000  
DEFINITION Mycobacterium tuberculosis H37Rv RNA-polymerase beta subunit (rpoB)  
gene, partial cds.  
ACCESSION U12205  
VERSION U12205.1 GI:515684  
KEYWORDS  
SOURCE Mycobacterium tuberculosis.  
ORGANISM Mycobacterium tuberculosis.  
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
Actinomycetales; Corynebacterineae; Mycobacteriaceae;  
Mycobacterium; Mycobacterium tuberculosis complex.  
1 (bases 1 to 3853)  
Imboden, P., Trollier, R., Marchesi, F., Telenti, A., Bodmer, T.,  
Cole, S., Schopfer, K. and Burkart, T.  
The rpoB gene of Mycobacterium tuberculosis  
unpublished  
2 (bases 1 to 3853)  
Imboden, P.  
Direct Submission  
Submitted (11-JUL-1994) Paul Imboden, Institute for Medical  
Microbiology, University of Berne, Friedbuehlstrasse 51, Berne,  
3010, Switzerland  
Location/Qualifiers  
1. .3853  
/organism="Mycobacterium tuberculosis"  
/strain="H37Rv"  
/db\_xref="taxon:1773"  
576. .>3853  
/gene="rpoB"  
576. .>3853  
/gene="rpoB"  
/codon\_start=1  
/transl\_table=11



KEYWORDS	Mycobacterium tuberculosis CDC1551.
SOURCE	Mycobacterium tuberculosis CDC1551
ORGANISM	Bacteria; Firmicutes; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium tuberculosis complex.
REFERENCE	1 (bases 1 to 19352)
AUTHORS	Fleischmann, R.D., Alland, D., Eisen, J.A., Carpenter, L., White, O., Peterson, J., DeBoy, R., Dodson, R., Gwinn, M., Haft, D., Hickey, E., Kolonay, J.F., Nelson, W.C., Umayam, L.A., Ermolaeva, M., Salzberg, S.L., Delcher, A., Utterback, T., Weidman, J., Khouri, H., Gill, J., Mikula, A. and Bishai, W.
TITLE	Whole genome comparison of Mycobacterium tuberculosis clinical and laboratory strains
JOURNAL	Unpublished
REFERENCES	2 (bases 1 to 19352)
AUTHORS	Fleischmann, R.D., Alland, D., Eisen, J.A., Carpenter, L., White, O., Peterson, J., DeBoy, R., Dodson, R., Gwinn, M., Haft, D., Hickey, E., Kolonay, J.F., Nelson, W.C., Umayam, L.A., Ermolaeva, M., Salzberg, S.L., Delcher, A., Utterback, T., Weidman, J., Khouri, H., Gill, J., Mikula, A. and Bishai, W.
TITLE	Direct Submission
JOURNAL	Submitted (25-APR-2001) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA
FEATURES	Location/Qualifiers
source	1. 19352 /organism="Mycobacterium tuberculosis CDC1551" /strain="CDC1551" /db_xref="taxon:83331" /note="clinical strain" 163..3699 /gene="MT0695" 163..3699 /gene="MT0695" /note="similar to GB:L27989 GB:L05910 GB:U12205 SP:P47766 PID:149992; identified by sequence similarity; putative" /codon_start=1 /transl_table=11 /product="DNA-directed RNA polymerase, beta subunit" /protein_id="AAK44921.1" /db_xref="GI:13880218"
gene	/translation="MLEGICLADSRQSKTAASPSRQSSNNSVPGAPNRVSFAKL REPLEVGLDQTDVLEWLGSPRMRSAERGVNVPVGLLEEVLYELSEDFSGS MSLSFSDPRDDKAPVDECKDKMTAAFLVTAEFINNNTGETIKSOTVPMGFPMX TEKGTFTLINTERVYVQLVSRPGVYFDETDKSTDKTLHSVKVIPSRGANLEFDVK RDVGVGRIDKRQPVTLKALGWTSQEIYVRFGESEIMRSTLEKONTVGTDEALLD IYKRLGEPPTKESAQTLLENLFFKAKRYDLARVRYKVKKKLGLHVGEPITSTLT EDVVAETYLRLHGGTMTVPGGVPEVPTDDIDHFNRLRLTVGELIQNQIRVG MSMRVYRMTQDVEAITPQTLINIPVVAIKEFFGTSQLSQFMDQNNPLSGLT HKRRLSALGPGSLRAGLEVDRVHPSHYGRMCPIETPEGNIGLSGLSVYARVNP EVEYVPSVEDVMDVSPROMSVATAMIFLEHDDANRLMGANNQORQAVFLVRSEAP LVCTGMELRAADAGDVVAESGVIEEVSADYITVHDNGTRRTYMRKFAFNHGT CANQCPIVADGRVEAGQVADPCTDDGEMALGNLLVAIMPWEGHNYEDAIIISNR LVVEDVLTSHIEEHDIDAKTKLGAETRIDPNI SDEVLADLDERIGVIRGAERD GTLVQVTPKGETLTPEERLLRAIFGKAREVRDTSIKVPHGESGKVGIRGVFSRE DEDELPAVGNELVRYVVAQRKISDQKLAGRKNKGVIKGLIPVEDMPTFLADCTPVD IILNTHGVPRNMIGQIILETHLNGCAHSCWKVDAAGVDPDWAARLPDELEAQNAIV STPVDSGAOEALGILLCTLPNRGQDVLVDAGDKAMLFDRSGSEPPYPVYVTMTLI MKLHLVDDKIHARSTGPGVSMITQPLGKAQFGQREGMECWAMQAYGAAYTQLEI LTKGDEDDLRVAVYEAIVKGENIPEGPISFVLLKELQSLCLNVEVLSSDGAATE LRGEDEDLRAAANLGINLSRNESASVEDLA"
gene	3744..7694 /gene="MT0696"
CDS	3744..7694 /gene="MT0696" /note="similar to SP:P37871; identified by sequence similarity; putative" /codon_start=1 /transl_table=11 /product="DNA-directed RNA polymerase, beta-prime subunit" /protein_id="AAK44922.1" /db_xref="GI:13880219"
gene	/translation="MLDVNFDELRLGLATAEDIRQMSYGEVKKPKETI NYRTLKPEKD GLFCEKIFGRDWECCYCKYKVRREKGIICERCYGEVTRAKVRERMGHIELAAPT HIYFKGVP SRLCYLLDAPKDLKLIYFAAYVITSVDEMRHNEHLSLEAMVERK AVEDORGELEAQAQLADLAELAEAGAKADARRKVRDEGGERMQINRAORELDR LEDWSTFTKLPKQLIVDENLYRELVDYRGSEYFTTGAMGAEISQKLIENFDIAEAS LRDVRNFKGQKKLRALKLVVAAPQSQSPMGVMDAVPVIPELRPMVQLDGGRR FATSDNLRYRVI NNRLKRLIDLCAPETI VNNKRMQLQESVDALFDNGRRGRPV GPNRPLKSLSDLLKGQFRONLLGKRDYSGRSVI VVGPKLKHQGLPKLMAL LKFPEVMLRLVDLHNAQNLKSAKMYERQRPQVMDVLEEVIAEPVILNARPTLRLGL IQAFPEMLVSGKAIQLHPLCYEAFNADFDQDNVHPLPSAQAQAEARTILMSSNNIL SPASGRPLAMPRLDMVTGLYLTVEPDETYEQPDGHPETGVYSSPAEAIMAADR GVLSVRAKIKRLLTQLRPPVEIEAEFGHSGQPGDMAETTLGRVNFNELLPLGYP FVAKQHKYQAAIINDLAERYPMI VVAOTVKLDAGFYATRGSTVSNADVLVPP RKKEILDHYBEADKVEKOFORGALNHDEALVEIKKEATDEVGQALREHYPDNP IITIVDSGATGNFOTRTLAGMKGLVTPKGEFIRPPKSSPREGTLVLEFINTHGA RKLADTALFADSGYLTTRLLVDVSDVIVREHDCQTERGIVVLAERAPDGTILRQ XIETSAITGLTADVADEAGNIVVERGQDQDPEIDALLAAGITQVKRSVLTCATST YCATCYGRSMATGKLVLDIGEAIVGAQSIGEPDQTLMTFTHOGGVGEDITGGLPR VQELFARVPRGRAPADVTGRVRLDEGRFYKITIPDDGGEEVYVYDKISKRQLRV FKHEDSERVLSGDHVEGQQLMEGSADPHVLRVQGPREVQIHLVREVEVYRAQG VSIHDKHIEVIRQMLRRVTIIDSGSTFLPGSLIDRAEFEEAENRRVVAEGEPAAGR PVLMTGKTASLATDSWLSAASFQETTRVLTDAAINCRSDKLNGLKENVIKGLIPAGT GINRYNIAVQPTPEERAAAYTIPSTYEDQYISPDFGAATGAAPVLDYDYSIDYR"
gene	complement(7691..8065) /gene="MT0697"
CDS	complement(7691..8065) /gene="MT0697" /note="identified by Glimmer2; putative" /codon_start=1 /transl_table=11 /product="hypothetical protein" /protein_id="AAK44923.1" /db_xref="GI:13880220"
gene	/translation="MFDSAALINPCHAWASAMERSGLLECAGLDEQPFGEFTADKL NPDGSSRRVPRQADGGTATVHERGGQSGGAGVVPQMHGFPALAMQDRLIHH GEQTQNRKIAQFNRVFCVCSPT"
gene	complement(8058..9972) /gene="MT0698"
CDS	/note="This region contains an authentic frame shift and is not the result of a sequencing artifact; identified by Glimmer2; putative; conserved hypothetical protein, authentic frameshift" 10167..10925 /gene="MT0699" 10167..10925 /gene="MT0699" /note="identified by match to PFAM protein family HMM pf01261" /codon_start=1 /transl_table=11 /product="AP endonuclease, family 2" /protein_id="AAK44924.1" /db_xref="GI:13880221"
gene	/translation="MLGSHVSPDPLAAAEAGADVQIFLGNPQSHKAPKPRDDAA AKKATLPIYHAPYLINLASANNRIPSRRKILQETCAAAADIGAAAVIVHGGHVAD DNDIDKGFQNRKALDTEVYVLENTAGDHAMARFDTIARLMDVIGDTGIGFC LDTCHTMAAGEALTADVRKAITGRIDLVHNCNDSRDEAGSRDRHANLAGSGDIDPL LVAAVKAAGAPVICETADQGRKDDIAFLRERTGS"
gene	10957..11799 /gene="MT0700"
CDS	10957..11799 /gene="MT0700" /note="similar to GB:U00012 PID:466863; identified by sequence similarity; putative" /codon_start=1 /transl_table=11 /product="hydrolase/esterase, putative" /protein_id="AAK44925.1" /db_xref="GI:13880222"
gene	/translation="MLRVAIILLAAVLAFAGCSGGTRLAAGFGNGNSVHTLDVGDAGR SYRLKPKVGLSPSAPLVYMLHGGFGSAQAKERSYGMDELADSEKFLVAYPDGYHRAWN ANGGCGCCGRAPAREVDDIGFVRAVVADIANNVSDPARVYVTGNSCAINSYTLACHT SIFAAGVYVSGTQDPCQSPRPVSVIHHTGTADPLVRYHGGPCAGFARIDGPPVDUN



AFWREVNRCALDITTEGPVTTSGATCADNRRVULLTVDDAGHRWPSFATQTLWRFFA  
AHRF"  
11859. .13487  
/gene="MT0701"  
11859. .13487  
/gene="MT0701"  
/note="similar to SP:P33224 GB:L20915 PID:457172  
PID:457174 PID:537028; identified by sequence similarity;  
putative"  
/codon\_start=1  
/transl\_table=11  
/product="acyl-CoA dehydrogenase, putative"  
/protein\_id="AAK44926.1"  
/db\_xref="GI:13880223"  
/translation="MSDTHVTNQVPPLENYNPASSPVLEIALIOEGGQWGLDEVNEV  
GATASCAQRWGELEADNRNPIILTHDGYRVYDEVEYDPAYHELMRTAITHGMHAAP  
WADRCAGHVHRAAKTSVMTVEPHICPTISMTVAVVPALRYNSLSAAVPEPLTSPREY  
DEPKPATKAGITACMSMTKEKOGSDVRAGTQATPNADGYSYLTGHKWFYSAPWCD  
IFVLQAAPDGLSCFLLPVLPDGTNRNMFLORLKDKLGNHANASSEVEYDGAVALV  
GEGRGVPIIENVNLTLDCAUGSATSMRTGLTRAVHQAHRKAFGAYLIDPOLMRN  
VLADLAVERAATIVAMRMAGATDNVRGNETEALLRRIGLAAAKYVWCKRSTAHAAE  
ALBCLNGYVEDSGPRPLYREAPLMGIWEGSNVSALDTRLAMATRPACVEVLFDEL  
ARSAGQDPRLDGHRVERLPQLGLDITIGYRARKIAEDICLALQGLSLVYRHRPAVAEA  
FLATRLGGOWGGAYGTMFAGLDLAPILERALYKG"  
13498. .14436  
/gene="MT0702"  
13498. .14436  
/gene="MT0702"  
/note="similar to GP:3885480; identified by sequence  
similarity; putative"  
/codon\_start=1  
/transl\_table=11  
/product="enoyl-CoA hydratase/isomerase family protein"  
/protein\_id="AAK44927.1"  
/db\_xref="GI:13880224"  
/translation="MTHAIRPVDNLTMTYEVTRGRIATFNRPKGNATIAIDTPL  
ELSAVERADLPDGVHIVILVSGRGEFCFADLSAYAGSSSTGGGAYGQVLDGKT  
QAVNHLPNPDMPIDYOMKSRFVRGFSALMHADRPVYKINGCYVAGGTDIALHADQ  
VIAAADAKGYPTPRVWGPVAGLWHRGLDQRAKRLFTGDCITGAQAAGLAVEA  
PEPADLTERLVAALPQLIMVKLALNSALLQOOGVATSRMVSTVFDDGAARHT  
PEHAFAVADAVEHGFDAVRRDEPFQDYGRQASRV"  
14439. .15161  
/gene="MT0703"  
14439. .15161  
/gene="MT0703"  
/note="identified by Glimmer2; putative"  
/codon\_start=1  
/transl\_table=11  
/product="hypothetical protein"  
Query Match 80.0%; Score 166.4; DB 1; Length 19352;  
Best Local Similarity 87.5%; Pred. No. 2.2e-18;  
Matches 182; Conservative 0; Mismatches 26; Indels 0; Gaps 0;  
QY 1 TCAAGGAGAGCGCTACGACCTGCCCGCGCTGGCGGCTACAGGTCAACAAGAGCTGG 60  
Db 1064 TCAAGGAGAGCGCTAGGACCTGCCCGCGCTGGCTATAGGTCACAGAAGCTCG 1123  
QY 61 GCCTGAACGCCGCCAGCCGATCAGCTGCTGACTCTGACCCGAGGAGACGTCGTGCCCA 120  
Db 1124 GGCTGATGTCGGCGAGCCCATCAGCTGCTGACGCTGACCGAAGAAGACGTCGTGCCCA 1183  
QY 121 CCATCGAGTACCTGGTGGCGCTCAGAGGCCGAGACCCAGGATGACCGTCCCGGGGGGG 180  
Db 1184 CCATCGAATATCTGGTGGCGCTTTCAGAGGGTCAGACCCAGGATGACCGTCCCGGGGGGG 1243  
QY 181 TCGAGGTCCCGGTCGAGGTCGACGACAT 208  
Db 1244 TCGAGGTCCCGGTGGAACCGACGACAT 1271  
RESULT 15  
MTCI376

LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
FEATURES  
source  
RBS  
gene  
CDS

MTCI376  
Mycobacterium tuberculosis H37Rv complete genome; segment 32/162.  
295972 AL123456  
295972.1 GI:3261790  
Mycobacterium tuberculosis H37Rv.  
Mycobacterium tuberculosis H37Rv  
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
Actinomycetales; Corynebacterineae; Mycobacteriaceae;  
Mycobacterium tuberculosis complex.  
1 (bases 1 to 19770)  
Cole, S.T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C.,  
Harris, D., Gordon, S.V., Eiglmeyer, K., Gas, S., Barry, III, C.E.,  
Tekala, F., Badcock, K., Basham, D., Brown, D., Chillingworth, T.,  
Connor, R., Davies, R., Devlin, K., Felwell, T., Gentles, S.,  
Hamlin, N., Holroyd, S., Hornsby, T., Jajelski, K., Krogh, A., McLean, J.,  
Mout, S., Murphy, L., Oliver, S., Osborne, J., Quail, M.A.,  
Rajandream, M.A., Rogers, J., Rutter, S., Seeger, K., Skelton, S.,  
Squares, S., Squares, R., Sulston, J.E., Taylor, K., Whitehead, S. and  
Barrell, B.G.  
Deciphering the biology of Mycobacterium tuberculosis from the  
complete genome sequence  
Nature 393 (6685), 537-544 (1998)  
98295987  
9634230  
2 (bases 1 to 19770)  
Parkhill, J.  
Direct Submission  
Submitted (11-JUN-1998) Submitted on behalf of the Mycobacterium  
tuberculosis sequencing and mapping teams, Sanger Centre, Wellcome  
Trust Genome Campus, Hinxton, Cambridge CB10 1SA Unite de Genetique  
Moleculaire Bacterienne, Institut Pasteur, 28 rue du Docteur Roux,  
75724 Paris Cedex 15, France E-mail: parkhill@sanger.ac.uk  
On Jun 27, 1998 this sequence version replaced gi:2143285.  
Notes:  
Details of M. tuberculosis sequencing at the Sanger Centre are  
available on the World Wide Web.  
(URL, http://www.sanger.ac.uk/projects/M\_tuberculosis/) CDS have  
been renumbered from the original cosmid submissions but the old  
gene designations are in brackets after the new gene numbers.  
Gene prediction was based on a Hidden Markov Model of TB genes  
implemented in TBparse (Krogh) supplemented with visual inspection  
of positional base preference in codons, especially where there is  
an increase in the observed/expected third position G + C.  
CAUTION: In some cases we may not have predicted the correct  
initiation codon. Where possible we choose an initiation codon  
(atg, gtg, or ttg) which is preceded by an upstream ribosome  
binding site sequence (optimally 5-13bp before the initiation  
codon). If this cannot be identified we choose the most upstream  
initiation codon.  
Location/Qualifiers  
1..19770  
/organism="Mycobacterium tuberculosis H37Rv"  
/strain="H37Rv"  
/db\_xref="taxon:83332"  
/clone="I376"  
57..61  
/note="aggga, possible rbs upstream of Rv0654"  
68..1573  
/gene="Rv0654"  
68..1573  
/gene="Rv0654"  
/note="Rv0654", (MTCI376.22), len: 501. unknown. FASTA  
score: Q53353 LIGNOSTILBENE-ALPHA, BETA-DIOXYGENASE (485  
aa) opt:280 z-score: 330.1 E(): 2.3e-11, (28.5% identity  
in 523 aa overlap). Also similar to M. tuberculosis protein  
MTCI21C12.07c (29.5% identity in 522 aa overlap)"  
/codon\_start=1  
/transl\_table=11  
/product="hypothetical protein Rv0654"  
/protein\_id="CA809380.1"  
/db\_xref="GI:2143307"  
/db\_xref="SPTREMBL:O06785"



Db 10938 TCGAGGTCCCGGTGGAACCGACGACAT 10965

Search completed: November 13, 2002, 01:27:11  
Job time : 660.723 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic acid - nucleic search, using sw model

Run on: November 12, 2002, 16:59:42 ; Search time 1109.13 Seconds  
(without alignments)  
3037.202 Million cell updates/sec

Title: US-09-697-123b-17  
Perfect score: 208  
Sequence: 1 tcaaggaggaagcgtacgac.....ccggtggaacccgacgacat 208

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estnu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vrt:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_nam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rod:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	155.2	74.6	1282	9	AI770311 42 Mycoba
2	48.2	23.2	649	12	BG908023 Talr1164F
3	48.2	23.2	1013	10	BE418320 SCL024.F0
C 4	48	23.1	487	12	BF444643 262617.MA
5	48	23.1	704	13	BM190539 POSM01000
6	46.6	22.4	363	12	BG907534 Talr1160G

RESULT 1  
AI770311/c  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
COMMENT

AI770311 1282 bp mRNA linear EST 24-JAN-2000  
42 Mycobacterium anaerobic stationary phase library Mycobacterium smegmatis CDNA, mRNA sequence.

AI770311 GI:6742680

EST.

Mycobacterium smegmatis.

Mycobacterium smegmatis

Bacteria: Actinobacteria; Actinobacteridae; Actinomycetales;

Corynebacterineae; Mycobacteriaceae; Mycobacterium.

1 (bases 1 to 1282)

Murugasu-Oei, B., Tay, A. and Dick, T.

Upregulation of stress response genes and ABC transporters in

anaerobic stationary-phase Mycobacterium smegmatis

Mol. Gen. Genet. 262 (4-5), 677-682 (1999)

20092472

Contact: Murugasu-Oei, B.

Mycobacterium Laboratory

Institute of Molecular and Cell Biology

30 Medical Drive, Singapore 117609, Republic of Singapore

Tel: 65 874 3011

Fax: 65 779 1117

Email: mcbom@mcmb.nus.edu.sg

Insert Length: 1282 Std Error: 0.00

Seq primer: T3 Forward; T7 Backward.

Location/Qualifiers

1..1282

source

BJ233830 BJ233830  
BJ258518 BJ258518  
BG349727 947030C07  
AL830817 AL830817  
BQ577676 3524\_1\_39  
BQ579756 WHE2974\_B  
AL830932 AL830932  
AL8056150 660004G09  
AL236849 Tetraodon  
BQ248115 Talr1156D  
BG908999 Talr1156D  
BQ246063 Talr15016F  
BM285614 524293.MA  
BE476577 159902.BA  
BE756092 210134.MA  
AL205895 Tetraodon  
BQ754171 EBCa01\_SQ  
BG369919 HVSME1002  
BV835256 AV835256  
BJ469527 BJ469527  
BI778452 EBR007.SQ  
BF618921 HVSMEC000  
BF624572 HVSMEa001  
AY107010 Zee mays  
AI676691 etnEST046  
AL831237 AL831237  
BG354527 947038E04  
AW564194 LG1\_285\_B  
BI722737 I031063H0  
BI960536 HVSMEH002  
BM135567 WHE2621\_D  
BJ267444 BJ267444  
AU070110 AU070110  
BG416431 HVSMEK001  
BI489596 603032031  
BE196164 HVSMEH009  
BJ470020 BJ470020  
AW484285 59867.MAR  
AU090131 AU090131

#### ALIGNMENTS

7 13 517 13 21.6 45  
8 13 684 13 21.6 45  
9 12 448 12 21.1 43.8  
10 9 468 9 21.1 43.8  
11 14 514 14 20.7 43.8  
12 14 648 14 20.3 42.2  
13 9 413 9 20.3 42.2  
14 10 645 10 20.2 42.2  
15 17 896 17 20.2 42.2  
16 14 573 14 20.1 41.8  
17 12 630 12 20.1 41.8  
18 13 649 13 20.0 41.6  
19 13 512 13 20.0 41.6  
20 10 514 10 20.0 41.6  
21 12 534 12 20.0 41.6  
22 9 966 9 20.0 41.6  
23 14 458 14 19.9 41.4  
24 12 496 12 19.9 41.4  
25 10 506 10 19.9 41.4  
26 13 592 13 19.9 41.4  
27 13 799 13 19.9 41.4  
28 12 799 12 19.9 41.4  
29 12 870 12 19.8 41.2  
30 11 1489 11 19.6 40.8  
31 9 462 9 19.5 40.6  
32 9 471 9 19.5 40.6  
33 12 485 12 19.5 40.6  
34 10 577 10 19.5 40.4  
35 13 616 13 19.4 40.4  
36 13 656 13 19.4 40.4  
37 13 512 13 19.3 40.2  
38 13 717 13 19.3 40.2  
39 9 460 9 19.2 40.0  
40 12 672 12 19.1 39.8  
41 13 704 13 19.1 39.8  
42 10 935 10 19.1 39.6  
43 13 408 13 19.0 39.6  
44 10 416 10 19.0 39.6  
45 9 420 9 19.0 39.6

```

/organism="Mycobacterium smegmatis"
/strain="mc2155"
/db_xref="taxon:1772"
/clone_lib="Mycobacterium anaerobic stationary phase library"
/lab_host="E. coli XL1-Blue MRF'"
/notes="Vector: Lambda ZAP II; Bacilli were disrupted using a Mini bead beater (Biospec). RNA was isolated using the RNeasy protocol (Qiagen). Purified RNA was subjected to 2 rounds of digestion with RNase-free DNase I (Promega). DNase I was heat-inactivated at 75°C for 5 min. and removed by using RNeasy columns followed by phenol extraction and ethanol precipitation. The RNA preparations were confirmed to be free of genomic DNA contamination by carrying out PCR and RT-PCR using the Access kit (Promega) and primers specific for the histone-like protein gene hlp (Lee et al., 1998). cDNA was synthesized using random hexamer primers (Promega) and Stratagene's cDNA synthesis kit. cDNA fragments were ligated into lambda ZAP II vector and packaged in vitro using Stratagene reagents."
BASE COUNT      230 a 395 c 405 g 252 t
ORIGIN

Query Match      74.6%; Score 155.2; DB 9; Length 1282;
Best Local Similarity 84.1%; Pred. No. 3.4e-25;
Matches 175; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 1 TCAAGGAGACCGGTAGACCTCGCGGGTGGCCGCTACAAGGTCAACAAGAGCTCG 60
Db 236 TCAAGGAGAACGCTAGACGCTGTGGCCGCTTACAAGGTCAACAAGAGCTCG 177
QY 61 GCCTGAACACCGCTCCCGCATACGACGACCACTCTGACCGAGAGAGCGTGTGCCCA 120
Db 176 GCCTGAACCGGGCAACCGCATCACAGCTCGACGCTGACCGAGAGAGCGTGTGCCGA 117
QY 121 CCATCGAGTACTGCTCGGCTGCACGAGGGCCACACGATGACCGTCCCGGGGGGAG 180
Db 116 CCATCGAGTACTGCTCGGCTGCACGAGGTCAGACTTCGATGACCGTCCCGGGTGGCG 57
QY 181 TCGAGGTGCGGTGGAAACGACGACAT 208
Db 56 TCGAGGTTCGCGTGGAGGTGCAGGACAT 29

RESULT 2
BG908023      649 bp mRNA linear EST 05-JUN-2001
LOCUS      TaLr1164F08r1 TaLr1 Triticum aestivum cDNA clone TaLr1164F05 5',
DEFINITION      mRNA sequence.
ACCESSION      BG908023
VERSION      BG908023.1 GI:14315686
KEYWORDS      EST.
SOURCE      bread wheat.
ORGANISM      Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae
; Triticeae; Triticum.
1 (bases 1 to 649)
Cloutier,S., Dong,G. and Walsh,A.
Wheat functional genomics - Thatcher Lr1 cDNA library
Contact: Dr. Sylvie Cloutier
Cereal Research Centre, Agriculture and Agri-food Canada
195 Dafoe Rd, Winnipeg, MB, Canada R3T 2M9
Tel: (204) 983-2340
Fax: (204) 983-4604
Email: scloutier@agr.ca
was cloned directionally, not all sequences generated with reverse
primer were from the 5' end (same with forward primer and 3' end).
Average inset size is >2.2 kb
Plate: 164 row: F column: 05
Seq primer: M13 Reverse.

```

## FEATURES

source

Location/Qualifiers

1..649

```

/organism="Triticum aestivum"
/cultivar="Thatcher Lr1"
/db_xref="taxon:4565"
/clone="TaLr1164F05"
/clone_lib="TaLr1"
/tissue_type="Leaf tissue"
/dev_stage="14 Days old"
/lab_host="E. coli XL0LR"
/notes="Vector: Lambda ZapII; mass excised in plasmid
vector PBK-CMV (Stratagene).; Site 1: EcoRI; Site 2: XhoI;
mRNA obtained from wheat NIL Thatcher Lr1 24 hours after
inoculation with leaf rust pathogen Puccinia triticina
race BBB carrying the avirulence gene AvrL."
BASE COUNT      162 a 178 c 179 g 130 t
ORIGIN

```

```

Query Match      23.2%; Score 48.2; DB 12; Length 649;
Best Local Similarity 57.7%; Pred. No. 0.28;
Matches 86; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 25 CGCGGTGGCGCTACAGGTCAACAAGAGCTCGGCTGAACACCGGTCCCGCATCA 84
Db 190 CGCGGTGGCGCTACAGGTCAACAAGAGCTCGGCTGAACACCGGTCCCGCATCA 249
QY 85 CGACGACCACTCTGACCGAAGAGGAGCGTGTGCCACCATCGAGTACCTGGTCCGCTGC 144
Db 250 AGATCGCCCAAGGACACCCCGCACGCGGCCACCTGCGCCATCAAGGTGCTCGACCCGCAACC 309
QY 145 AGAGGGGCCACACCGATGACCGTCCCG 173
Db 310 ACGTCTCGGCCACAAGATGGTGTGAGCAG 338

```

## RESULT 3

BG418320

LOCUS

```

DEFINITION      BE418320      1013 bp mRNA linear EST 24-JUL-2000
                  SCL024.F08R990724 ITEC SCL Wheat Leaf Library Triticum aestivum
                  cDNA clone SCL024.F08, mRNA sequence.
ACCESSION      BE418320
VERSION      BE418320
KEYWORDS      EST.
SOURCE      bread wheat.
ORGANISM      Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae
; Triticeae; Triticum.
1 (bases 1 to 1013)
Anderson,O.A., Appels,R., Bailey,P., Blake,T., Close,T., Cloutier
,S., Dubcovsky,J., Feuillet,C., Gale,M., Graner,A., Gustafson,P.,
Herrmann,R.G., Holton,T., Jacquemin,J.M., Jia,J., Joudrier,P.,
Langridge,P., Lazo,G.R., Lin,J.J., McGuire,P., Ogihara,Y.,
Pecchioni,N., Qualset,C., Schuch,W., Selvaraj,G., Shariflou,M.,
Sorrells,M., Warburton,M. and Wenzel,G.
International Triticeae EST Cooperative (ITEC): Production of
Expressed Sequence Tags for Species of the Triticeae
Unpublished (2000)
Contact: Cloutier S
Cereal Research Centre, Agriculture & Agri-Foods Canada
Winnipeg MT CANADA
Tel: 204 983 2340
Fax: 204 983 4604
Email: scloutier@agr.ca
International Triticeae EST Cooperative (ITEC)
http://wheat.pw.usda.gov/genome.

```

## TITLE

JOURNAL

COMMENT

## FEATURES

source

Location/Qualifiers

1..1013

```

/organism="Triticum aestivum"
/cultivar="Thatcher Lr1"
/db_xref="taxon:4565"
/clone="SCL024.F08"
/clone_lib="ITEC SCL Wheat Leaf Library"

```

```

/tissue_type="young leaf"
/dev_stage="after 24 hour challenge with LR-AVR1"
/note="Vector: Lambda ZAP; i.0 Kbp average insert size."
BASE COUNT      246 a      237 g      223 t      36 others
ORIGIN

Query Match      23.2%; Score 48.2; DB 10; Length 1013;
Best Local Similarity 57.7%; Pred. No. 0.28; 63; Indels 0; Gaps 0;
Matches 86; Conservative 0; Mismatches 0;

QY 25 CCGGGGTGGCGGTACAAAGTCAACAAGAGCTCGGCTGAACACCGCGTCCCGCATCA 84
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 186 CCGGGTGGCGGCTACGAGTTCGCAAGACCGTTCGGGAGGGGAGCTTCCCGAAGTCA 245
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 85 CGACGACCACTCTGACCAAGAGAGCTGCTGCGGCACATCGAGTATGCTGCGCCTGC 144
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 246 AGATCGCAAGAGACACCGCAACCGCCACCTCGGCATCAAGGTGCTGCGACGCAACC 305
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 145 ACGAGGCGCCACACGACGATGACCTGCCG 173
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 306 ACGTCTCCGCCACANATGTCGAGCAG 334
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

RESULT 4
BF444643/c
LOCUS      BF444643      487 bp      mRNA      linear      EST 01-DEC-2000
DEFINITION 26217 MARC 2P1G Sus scrofa cDNA 5', mRNA sequence.
ACCESSION  BF444643
VERSION    BF444643.1 GI:11504735
KEYWORDS   EST.
SOURCE     pig.
ORGANISM   Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE  1 (bases 1 to 487)
AUTHORS   Fahrnkruug,S.C., Freking,B.A., Rohrer,G.A., Smith,T.P.L., Casas,E.,
            Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A., Laegreid,W.W.
            and Keele,J.W.
TITLE     Design and use of two pooled tissue normalized cDNA libraries for
            EST discovery in swine
JOURNAL   Unpublished (2000)
COMMENT   Contact: Smith TPL
            USDA, ARS, US Meat Animal Research Center
            PO Box 166, Clay Center, NE 68933-0166, USA
            Tel: 402 762 4366
            Fax: 402 762 4390
            Email: smith@mail.marc.usda.gov
            Single pass sequencing. Bases called and alt trimmed with phred
            v0.980904.e. Vector identified by cross_match with the -minscore 18
            and -minmatch 12 options.
PCR primers
FORWARD: AGCAACACGCTATGACCAT
BACKWARD: GTTTTCCAGTCACGACG
Plate: 98 row: L column: 4
Seq primer: ATTTAGTGCACACTATAG.
FEATURES
            source
            1..487
            /organism="Sus scrofa"
            /db_xref="taxon:9823"
            /clone_lib="MARC 2P1G"
            /tissue_type="pooled"
            /lab_host="DH10B"
            /note="Vector: pCMV SPORT6; Site_1: NotI; Site_2: SalI;
            Library made from pooled tissue from testis, ovary,
            endometrium, hypothalamus, pituitary, and placenta."
BASE COUNT      72 a      141 c      162 g      112 t
ORIGIN

Query Match      23.1%; Score 48; DB 12; Length 487;
Best Local Similarity 51.9%; Pred. No. 0.3;
Matches 108; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

QY 1 TCAAGGAGAAGCGCTACGACCTCGCGGGTGGCGGCTTACAAAGTCAACAAGAGCTCG 60
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

```

```

|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 408 TCAAGGAAGCCATCAAGAAGCTGCAGTGGATGGCGGAGGCACGTTACAGGGGCGAGCCCC 349
QY 61 GCCTGAACACCGCGTCCCGATCACAGCAGCACTCTTGACCGAAGAGAGAGCTGTCGCCCA 120
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 348 TGCAGTACACCGCGGAGCGAGCTGCTGCCCGCCACTCAGAACAACCGGATCGCCCTGGTCA 289
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 121 CCATCGAGTACCTGGTCCCGCTGCACGAGGCCACACACGATGACCGTCCCGGGCGGAG 180
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 288 TCAGGAGCGCCCTCAGACACCCAGAGGACACACCCCGCTGAGCGTGTCTCGCGGCC 229
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 181 TCGAGTGGCGGTGGAACCGAGCAGCAT 208
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 228 CCGACATCCAGTGGTCTCGGTGGGCAT 201
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

RESULT 5
BM190539
LOCUS      BM190539      704 bp      mRNA      linear      EST 11-DEC-2001
DEFINITION POSM0100013_D07F porcine skeletal muscle cDNA library (PoSM) Sus
            scrofa cDNA 5', mRNA sequence.
ACCESSION  BM190539
VERSION    BM190539.1 GI:17526502
KEYWORDS   EST.
SOURCE     pig.
ORGANISM   Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE  1 (bases 1 to 704)
AUTHORS   Yao,J.,, Cousens,P., Ernst,C. and Saama,P.
            Analysis of expressed sequence tags from a normalized porcine
            skeletal muscle cDNA library
JOURNAL   Unpublished (2002)
COMMENT   Contact: Jianbo Yao
            Department of Animal Science
            Michigan State University
            B215 Anthony Hall, East Lansing, MI, USA
            Tel: 517-355-8443
            Fax: 517-353-1699
            Email: yaoj@msu.edu
            Seq primer: M13 reverse.
            Location/Qualifiers
            1..704
            /organism="Sus scrofa"
            /db_xref="taxon:9823"
            /clone_lib="porcine skeletal muscle cDNA library (PoSM)"
            /sex="male and female"
            /tissue_type="skeletal muscles"
            /dev_stage="45 d and 90 d of gestation, birth, 7 wk and 1
            yr of age"
            /lab_host="DH10B"
            /note="Organ: hind limbs; Vector: pSPORT1; Site_1: NotI;
            Site_2: SalI"
BASE COUNT      138 a      253 c      207 g      106 t
ORIGIN

Query Match      23.1%; Score 48; DB 13; Length 704;
Best Local Similarity 51.9%; Pred. No. 0.31;
Matches 108; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

QY 1 TCAAGGAGAAGCGCTACGACCTCGCGGGTGGCGGCTTACAAAGTCAACAAGAGCTCG 60
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 14 TCAAGGAAGCCATCAAGAAGCTGCAGTGGATGGCGGAGGCACGTTACAGGGGCGAGCCCC 73
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 61 GCCTGAACACCGGTCCCGATCACAGCAGCACCTCTGACCGAAGAGAGAGCTGTCGCCCA 120
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 74 TGCAGTACACCGGAGCGAGCTGCTGCCCGCCACTCAGAACAACCGGATCGCCCTGGTCA 133
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 121 CCATCGAGTACCTGTCGGCTGCACGAGGCCACACACGATGACCGTCCCGGGCGGAG 180
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 134 TCAGGAGCGCGCTCAGACACCCAGAGGACACACCCCGCTGAGCGTGTCTCGCGGCC 193
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 181 TCGAGGTGGCGGTGGAAACCGAGCAGCAT 208
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

```

```

Db      194 CCGACATCCAGGTGGTCTTCGTTGGGCAT 221
||||| | | ||||| ||| | |||
RESULT 6
BG907534          363 bp mRNA linear EST 05-JUN-2001
LOCUS            TaLr1l160G09R TaLr1 Triticum aestivum cDNA clone TaLr1160G09 5',
DEFINITION      mRNA sequence.
ACCESSION       BG907534
VERSION         BG907534.1 GI:14315210
KEYWORDS        EST.
SOURCE          bread wheat.
ORGANISM        Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
; Triticeae; Triticum.
REFERENCE       1 (bases 1 to 363)
AUTHORS        Cloutier,S.; Dong,G. and Walsh,A.
TITLE          Wheat functional genomics - Thatcher Lr1 cDNA library
COMMENT        Unpublished (2001)
Contact: Dr. Sylvie Cloutier
Cereal Research Centre, Agriculture and Agri-food Canada
195 Dafoe Rd, Winnipeg, MB, Canada R3T 2M9
Tel: (204) 983-2340
Fax: (204) 983-4604
Email: scloutier@agr.ca
was cloned directionally, not all sequences generated with reverse
primer were from the 5' end (same with forward primer and 3' end).
Average inset size is >2.2 kb
Plate: 160 row: G column: 09
Seq primer: M13 Reverse.
FEATURES             Location/Qualifiers
     source           1..363
     organism="Triticum aestivum"
     cultivar="Thatcher Lr1"
     db_xref="taxon:4565"
     /clone_lib="TaLr1l160G09"
     /tissue_type="Leaf tissue"
     /dev_stage="14 Days old"
     /lab_host="E. coli XL0LR"
     /note="Vector: Lambda ZapII; mass excised in plasmid
vector PBK-CMV (Stratagene); Site_1: EcoRI; Site_2: XhoI;
mRNA obtained from wheat NIL Thatcher Lr1 24 hours after
inoculation with leaf rust pathogen Puccinia triticina
race BBB carrying the avirulence gene Avr1."
BASE COUNT          90 a 112 c 107 g 54 t
ORIGIN
Query Match                22.4%; Score 46.6; DB 12; Length 363;
Best Local Similarity      57.0%; Pred. NO. 0.62;
Matches 85; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

Qy    25 CGCGGGTGCGCCGCTACAAGGTCAACAAGAAGCTCGGCCCTGAACACCGCGTCCCGCATCA 84
||||| ||||| ||| | ||| ||||| ||| | ||| ||||| ||| | |||
Db    176 CGCGGGTGCGCCGCTACGAGCTCGGACAGACCGCTCGGCGAGGCGAGCTTCTCCAAGGTCA 235
||||| ||||| ||| | ||| ||||| ||| | ||| ||||| ||| | |||
Qy    85 CGAGGACCCTCTGACCGAAGAGGACGTGTCGCCACCATCGAGTACTGTCGCCCTGC 144
||| ||||| ||| | ||| ||||| ||| | ||| ||||| ||| | |||
Db    236 AGATCGCCCAAGGACACCGCGCACCTCGGCCATCAAGTGTCTCGACCGCAACC 295
||| ||||| ||| | ||| ||||| ||| | ||| ||||| ||| | |||
Qy    145 ACAGAGGCCACACCGATGACCGTCCCCG 173
||| ||||| ||| | ||| ||||| ||| | ||| ||||| ||| | |||
Db    296 ACGTCTCCGGCACAAAGATGGTCGAGCAG 324
||| ||||| ||| | ||| ||||| ||| | ||| ||||| ||| | |||

RESULT 7
BJ233830
LOCUS            BJ233830          517 bp mRNA linear EST 05-APR-2002
DEFINITION      BJ233830 Y. Ogihara unpublished cDNA library, Wh_e Triticum
aestivum cDNA clone whe7g20 5', mRNA sequence.
ACCESSION       BJ233830

```







```

Db 196 GCGGCGCGCGGGAGGTCTCCACACGAGGCGCGCGCTCGGTACGGGCCCGGGA 255
QY 88 CGACCACTGTGACCGAAGAGGAGTGTGCGCCACCATCGAGTACCTGTCGCCCTGCACG 147
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 256 CGATGGCGCTGTGAGGTTGCGGATGTCGGGCGCCATCGCTACCTGTGTGTACCAGA 315
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 148 AGGCCACACACCATGACGCTCCCGGGCGGAGTCGAGTGCCTGCGTAACACCGAGAC 206
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 316 AGCGAGGCCCGCGCGCGCGGACGAGGTGCGCAAGGTGCGCGTCGGACACGCGGAC 374

RESULT 13
AL830932
LOCUS
DEFINITION AL830932 413 bp mRNA linear EST 16-JUL-2002
SEQUENCE.
ACCESSION AL830932
VERSION AL830932.1 GI:21842712
KEYWORDS EST.
SOURCE bread wheat.
ORGANISM
Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
1 (bases 1 to 413)
AUTHORS Wilson,I., Beswick,R., Shepherd,S., Barker,G., Parker,J., Owen,P.,
Edwards,D., Coghill,J., Holdsworth,M., Lenton,J., Shewry,P. and
Edwards,K.
TITLE A BSRG-funded wheat EST resource for the academic community
JOURNAL Unpublished (2002)
COMMENT Contact: Barker G
Institute of Arable Crop Research
Long Ashton, Bristol BS41 9AF United Kingdom.
FEATURES
Source
1..413
/organism="Triticum aestivum"
/cultivar="florida"
/db_xref="taxon:4565"
/clone="C04_q242_plate_15"
/clone_lib="q:242"
/tissue_type="egg cell"
/dev_stage="pre-fertilisation"
BASE COUNT 83 a 129 c 123 g 77 t 1 others
ORIGIN

Query Match 20.3%; Score 42.2; DB 9; Length 413;
Best Local Similarity 51.9%; Pred. No. 6;
Matches 95; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 26 GCGGTGGCGCGCTACAAAGGTCAACAAGCTCGGCTGAACACCGCGTCCCGGATCAC 85
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 106 GGGGCTGTCACCTACATTGTTCAGGACGGGCTGGAGTGACCCCATGTCCGCCATCTC 165
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 86 GACGACCACTGTGACCGAAGAGAGCTGTGCGCCACCATCGAGTACCTGGTCCGCTGCA 145
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 166 CAGCATCACCTCTCATCAACAAGTTTCAGCGTCGACAGAGCGTGGAGTCGCCGAGAAGTT 225
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 146 CGAGGCGCACACCATGACGTCGCCGGGAGTGCAGGTGCGCGTGGAACCCACGA 205
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 226 CGTCAGCTGGCGATGACGAGGGGCTCGGCTCCTCGAGTGGCAGTGGCTCCGACAC 285
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 206 CAT 208
    |||
Db 286 CGT 288

RESULT 14
AW056150
LOCUS
DEFINITION AW056150 645 bp mRNA linear EST 27-SEP-1999
SEQUENCE.
ACCESSION AW056150
VERSION AW056150.1 GI:5928858
KEYWORDS mRNA sequence.
SOURCE Human gene number estimate provided by genome wide analysis using
Tetraodon nigroviridis DNA sequence

```

```

KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACG
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 645)
AUTHORS Walbot,V.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford
University
JOURNAL Unpublished (1999)
COMMENT Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 660004 row: G column: 09.
FEATURES
Location/Qualifiers
1..645
/organism="Zea mays"
/cultivar="Ohio43"
/db_xref="taxon:4577"
/clone_lib="660 - Mixed stages of anther and pollen"
/tissue_type="whole premeiotic anthers to pollen shed"
/dev_stage="premeiotic anthers to pollen shed"
/lab_host="XL0LR"
/note="Organ: anthers; Vector: Lambda Zap; Site_1: EcoRI;
Site_2: XhoI; Anther and pollen cDNA library.
Directionally sequenced with 5' end at the EcoRI site.
Created by Amie Franklin."
BASE COUNT 121 a 200 c 221 g 103 t
ORIGIN

Query Match 20.3%; Score 42.2; DB 10; Length 645;
Best Local Similarity 57.0%; Pred. No. 6.1;
Matches 77; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 58 TCGGCTGAACACGCGTCCCGCATCAGCAGCACCTCTGACCGAAGAGAGCGTGTGTCG 117
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 53 TCGGCTGGCATCTGGGCGCGGAGTCCGACGTCCTGACGACGACGAGCGTGTGTCG 112
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 118 CCACCATCGATCTGTGTCGCTGCAGGAGGCCACACGATGACCGTCCCGGGCG 177
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 113 GCACATTCTGGTACTGGCGCGGAGTACTTCATGTACGGAGGTGACCGACAGCGTGTG 172
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 178 GAGTCGAGTCCCG 192
    ||| ||| ||| |||
Db 173 ACGTACGCGTTGG 187
    ||| ||| ||| |||

RESULT 15
CNS03BQW/C
LOCUS
DEFINITION CNS03BQW 896 bp DNA linear GSS 15-MAY-2000
SEQUENCE.
Tetraodon nigroviridis genome survey sequence PUC-Ori end of clone
012P03 of library G from Tetraodon nigroviridis, genomic survey
sequence.
AL236849
VERSION AL236849.1 GI:7895984
KEYWORDS GSS: genome survey sequence.
SOURCE Tetraodon nigroviridis.
ORGANISM Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodon.
1 (bases 1 to 896)
AUTHORS Roest-Crolius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,
Saurin,W. and Weissenbach,J.
TITLE Human gene number estimate provided by genome wide analysis using
Tetraodon nigroviridis DNA sequence

```



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 12, 2002, 16:05:40 : Search time 21.1607 Seconds  
(without alignments)  
3487.380 Million cell updates/sec

Title: US-09-697-123b-17

Perfect score: 208

Sequence: 1 tcaagagagacgctacgac.....ccggtggaacacgacgacat 208

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 320260 seqs, 177392727 residues

Total number of hits satisfying chosen parameters: 640520

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published\_Applications\_NA.\*  
1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/2/pubpna/US05\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*  
10: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*  
12: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq.\*  
13: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*  
14: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	81.6	39.2	5096	10 US-09-984-711-5	Sequence 5, Appli
2	81.6	39.2	5099	9 US-10-075-460-5	Sequence 5, Appli
3	81.6	39.2	5099	10 US-09-887-052-1	Sequence 1, Appli
4	81.6	39.2	5099	10 US-09-887-052-3	Sequence 3, Appli
5	81.6	39.2	5099	10 US-09-887-052-5	Sequence 5, Appli
6	38	18.3	720	9 US-09-999-745-14	Sequence 14, Appli
7	37	17.8	4039	10 US-09-205-448-7	Sequence 7, Appli
8	36.8	17.7	1182	10 US-09-799-514-7	Sequence 7, Appli
9	36.4	17.5	315	10 US-09-923-876-3629	Sequence 3629, Ap
10	36.4	17.5	394	10 US-09-960-352-9522	Sequence 9522, Ap
11	36.4	17.5	1665	10 US-09-815-242-7935	Sequence 7935, Ap
12	36.4	17.5	2329	10 US-09-816-828-9	Sequence 9, Appli
13	36.2	17.4	1704	10 US-09-815-242-7788	Sequence 7788, Ap
14	36.2	17.4	3300	10 US-09-379-931-6	Sequence 6, Appli
15	35.8	17.2	639	10 US-09-815-242-7939	Sequence 7939, Ap
16	35.8	17.2	2163	10 US-09-939-408A-29	Sequence 29, Appli
17	35.8	17.2	13842	10 US-09-861-289-30	Sequence 30, Appli
18	35.8	17.2	36778	10 US-09-861-289-5	Sequence 5, Appli
19	35.4	17.0	1206	10 US-09-815-242-7692	Sequence 7692, Ap

Sequence 7, Appli  
Sequence 19, Appli  
Sequence 2, Appli  
Sequence 2, Appli  
Sequence 2, Appli  
Sequence 2, Appli  
Sequence 2, Appli  
Sequence 2, Appli  
Sequence 7738, Ap  
Sequence 20, Appli  
Sequence 7696, Ap  
Sequence 7759, Ap  
Sequence 2599, Ap  
Sequence 13270, A  
Sequence 12569, A  
Sequence 1176, Ap  
Sequence 13975, A  
Sequence 13975, A  
Sequence 13975, A  
Sequence 1651, Ap  
Sequence 10408, A  
Sequence 4774, Ap  
Sequence 12, Appli  
Sequence 13, Appli  
Sequence 15, Appli  
Sequence 64, Appli  
Sequence 1, Appli  
Sequence 7925, Ap

US-09-984-711-5  
; Sequence 5, Application US/09984711  
; Patent No. US20020119549A1  
; GENERAL INFORMATION:  
; APPLICANT: MOECKEL, Bettina  
; APPLICANT: BATHE, Brigitte  
; APPLICANT: STEPHAN, Hans  
; APPLICANT: KREUTZER, Caroline  
; APPLICANT: HERMANN, Thomas  
; APPLICANT: PFEFFERLE, Walter  
; APPLICANT: BINDER, Michael  
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES WHICH CODE FOR THE rpsL GENE  
; FILE REFERENCE: 204209USO  
; CURRENT APPLICATION NUMBER: US/09/984.711  
; CURRENT FILING DATE: 2001-10-31  
; PRIOR APPLICATION NUMBER: DE10108230.9  
; PRIOR FILING DATE: 2001-02-16  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 5  
; LENGTH: 5096  
; TYPE: DNA  
; ORGANISM: Corynebacterium glutamicum  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (702)..(4196)  
; OTHER INFORMATION:  
US-09-984-711-5

#### ALIGNMENTS

RESULT 1  
US-09-984-711-5  
; Sequence 5, Application US/09984711  
; Patent No. US20020119549A1  
; GENERAL INFORMATION:  
; APPLICANT: MOECKEL, Bettina  
; APPLICANT: BATHE, Brigitte  
; APPLICANT: STEPHAN, Hans  
; APPLICANT: KREUTZER, Caroline  
; APPLICANT: HERMANN, Thomas  
; APPLICANT: PFEFFERLE, Walter  
; APPLICANT: BINDER, Michael  
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES WHICH CODE FOR THE rpsL GENE  
; FILE REFERENCE: 204209USO  
; CURRENT APPLICATION NUMBER: US/09/984.711  
; CURRENT FILING DATE: 2001-10-31  
; PRIOR APPLICATION NUMBER: DE10108230.9  
; PRIOR FILING DATE: 2001-02-16  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 5  
; LENGTH: 5096  
; TYPE: DNA  
; ORGANISM: Corynebacterium glutamicum  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (702)..(4196)  
; OTHER INFORMATION:  
US-09-984-711-5

Query Match 39.2%; Score 81.6; DB 10; Length 5096;  
Best Local Similarity 66.5%; Pred. No. 1.6e-12;  
Matches 133; Conservative 0; Mismatches 64; Indels 3; Gaps 1;  
Qy 9 AAGCGCTACGACCTCGCGCGGTGGCGCTACAAGGTCAACAAGAGCTCGCGCTGAAC 68  
|||||  
Db 1578 AAGCGCTACGACCTCGCGGTGGCGCTACAAGGTCAACAAGATCAACCGCAGCTCGCGCTGGT 1637  
Qy 69 ACCCGTCCCGATCAGCAGGACCACTTCAACGAGAGGACGCTCGTGGCGCACCATCGAG 128  
|||  
Db 1638 GGCACCAACGATGTTTGTATG---ACTCTTACTGAAGAGGACATCGCAACCAACCATCGAG 1694



Db 1638 GCGCACCAGCATGTTGATG---ACTCTTACTGAAGAGGACATCGCAACCAACCATCGAG 1694  
QY 129 TACCTGTGGCTGGCGTGCACGAGGGCCACACAGCATGACCGCTCCGGCGGAGTCGAGGTG 188  
Db 1695 TACCTGTGGCTGGCGTGCACGAGGGCCACACAGCATGACCGCTCCGGCGGAGTCGAGGTG 188  
QY 189 CCGGTGGAACCGAGCAT 208  
Db 1755 CCAGTCGAGACCGATGACAT 1774

## RESULT 5

US-09-887-052-5  
; Sequence 5, Application US/09887052  
; Patent No. US20020119537A1  
; GENERAL INFORMATION:  
; APPLICANT: MOECKEL, Bettina  
; APPLICANT: BATHE, Brigitte  
; APPLICANT: HERMANN, Thomas  
; APPLICANT: PFEFFERLE, Walter  
; APPLICANT: BINDER, Michael  
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES WHICH CODE FOR THE rpoB GENE  
; FILE REFERENCE: 204212US0X  
; CURRENT APPLICATION NUMBER: US/09/887,052  
; CURRENT FILING DATE: 2001-06-25  
; PRIOR APPLICATION NUMBER: DE10107229.5  
; PRIOR FILING DATE: 2001-02-16  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 5  
; LENGTH: 5099  
; TYPE: DNA  
; ORGANISM: Corynebacterium glutamicum  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (702)..(4196)  
US-09-887-052-5

Query Match 39.2%; Score 81.6; DB 10; Length 5099;  
Best Local Similarity 66.5%; Pred. No. 1.6e-12;  
Matches 133; Conservative 0; Mismatches 64; Indels 3; Gaps 1;  
QY 9 AAGCGCTACGACCTCGCGCGGTGGCGGTACAAAGTCAACAAAGCTCGCGCTGAAC 68  
Db 1578 AAGCGCTACGACCTCGCGGTGGCGGTACAAAGTCAACAAAGCTCGCGCTGAAC 68  
QY 69 ACCGCTCCCGATCAGCAGCACCCTCTGACCGAAGAGACGTCGTCGCCACCATCGAG 128  
Db 1638 GCGCACCAGATGTTTATG---ACTCTTACTGAAGAGGACATCGCAACCAACCATCGAG 1694  
QY 129 TACCTGTGGCTGGCGTGCACGAGGGCCACACAGCATGACCGCTCCGGCGGAGTCGAGGTG 188  
Db 1695 TACCTGTGGCTGGCGTGCACGAGGGCCACACAGCATGACCGCTCCGGCGGAGTCGAGGTG 188  
QY 189 CCGGTGGAACCGAGCAT 208  
Db 1755 CCAGTCGAGACCGATGACAT 1774

## RESULT 6

US-09-999-745-14  
; Sequence 14, Application US/09999745  
; Patent No. US20020157120A1  
; GENERAL INFORMATION:  
; APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA  
; APPLICANT: Tsien, Roger Y.  
; APPLICANT: Baird, Geoffrey  
; TITLE OF INVENTION: CIRCULARLY PERMUTED FLUORESCENT PROTEIN INDICATORS  
; FILE REFERENCE: REGEN1470-1  
; CURRENT APPLICATION NUMBER: US/09/999,745  
; CURRENT FILING DATE: 2001-10-23  
; PRIOR APPLICATION NUMBER: 09/316,920

; PRIOR FILING DATE: 1999-05-21  
; NUMBER OF SEQ ID NOS: 67  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 14  
; LENGTH: 720  
; TYPE: DNA  
; ORGANISM: Aequorea victoria  
; US-09-999-745-14

Query Match 18.3%; Score 38; DB 9; Length 720;  
Best Local Similarity 49.0%; Pred. No. 0.16;  
Matches 101; Conservative 0; Mismatches 105; Indels 0; Gaps 0;  
QY 1 TCAAGGAGAAGCGCTTACGACCTCGCGGGTGGCGGTACAAAGTCAACAAAGAGTCG 60  
Db 392 TCAAGGAGAAGCGCTTACGACCTCGCGGGTGGCGGTACAAAGTCAACAAAGAGTCG 60  
QY 61 GCCTGAACACCGCTCCCGGATCAGCAGGACCACTCTGACCGAAGAGAGTCGTCGCCA 120  
Db 452 TCTATATCACCCCGCACAAAGCAGAACGCGCATCAAGGCCCACTTCAAGATCCGCCACA 511  
QY 121 CCATCGACTACTGTGCTCGGCTGCAGGAGGCGCCACACAGCATGACCGTCCCGGGCGGAG 180  
Db 512 ACATCGAGGCGGAGCGTCGAGCTCGCGGACCACTACCAAGCAACACCCCATCGCGG 571  
QY 181 TCGAGGTCCCGGTGGAACCGACGAC 206  
Db 572 ACGGCGCGGTGCTGCTGCCGCAAC 597

## RESULT 7

US-09-205-448-7  
; Sequence 7, Application US/09205448  
; Patent No. US20020137661A1  
; GENERAL INFORMATION:  
; APPLICANT: Arnold, Frances  
; APPLICANT: Shao, Zhixin  
; APPLICANT: Volkov, Alexander  
; APPLICANT: California Institute of Technology  
; TITLE OF INVENTION: Method for Creating Polynucleotide and Polypeptide  
; FILE REFERENCE: 018097-025710PC  
; CURRENT APPLICATION NUMBER: US/09/205,448  
; CURRENT FILING DATE: 1998-12-04  
; PRIOR APPLICATION NUMBER: US 60/067,908  
; PRIOR FILING DATE: 1997-12-08  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 7  
; LENGTH: 4039  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:Actinoplanes  
; OTHER INFORMATION: utahensis echinocandin B (ECB) deacylase gene  
; OTHER INFORMATION: mutant M-15 created by in vitro heteroduplex  
; NAME/KEY: CDS  
; LOCATION: (1196)..(3559)  
US-09-205-448-7

Query Match 17.8%; Score 37; DB 10; Length 4039;  
Best Local Similarity 52.2%; Pred. No. 0.33;  
Matches 82; Conservative 0; Mismatches 75; Indels 0; Gaps 0;  
QY 20 CCTGCGCGGTGGCGGTGCTTACAAAGTCAACAAAGTCTGGCTGAACACCGCTGCC 79  
Db 2050 CATCGGATCGGCGCACAAACCGCAGCGTCGCTGGAGCCACACCGCTTCCACCGCCCGCGG 2109  
QY 80 GATCAGCAGCACCCTCTGTGACCGAAGAGGAGTCGTCGCCACCATCGAGTACCTGGTCCG 139  
Db 2110 GTTCGTGTGCGCACCGCTGTAGCCTCTGTCGCCGCGGACCCACCTCTCTATTAGTCGACGG 2169



```
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7935
; LENGTH: 1665
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1665)
US-09-815-242-7935

Query Match      17.5%; Score 36.4; DB 10; Length 1665;
Best Local Similarity 49.5%; Pred. No. 0.43;
Matches 94; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

QY 13 GCTACGACCTCCGCGGTGGCGGCTACAAAGGTCAACAAGAGTCTGGCGCTGAACACCG 72
Db 182 GCGAGGCGCGCGGCTGCTGCGGGGATCAACGCTCGGTACCTCCGCGAGGAGCCCAAGCTGG 241
QY 73 COTCCCCGATCATGACGACCACTCTGACCGAAGAGGAGTGTGTCGCCACCATCGAGTACC 132
Db 242 ACCCGAGCGCACCGTGGCGGCACATCGTCGAAGAGCGCGTGGCCAGATCAAGCAGGCCCC 301
QY 133 TGGTCCGCTCGACAGAGGGCCACACAGATGACCGTCCCGGGCGGAGTCGAGGTGCCGG 192
Db 302 AGGCGCGCTGGACGAGGTCTATGCGCGCTATGCCGAGCGGATGCCGACTTCGACGCCCC 361
QY 193 TGGAAACCGA 202
Db 362 TGGCGGCCGA 371

RESULT 12
US-09-816-828-9
; Sequence 9, Application US/09816828
; Patent No. US20020150898A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Goodrich, Ryle
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Xue, Aidong J.
; APPLICANT: Ma, Yungqing
; APPLICANT: Wang, Zhiwei
; APPLICANT: Zhao, Qing A.
; APPLICANT: Zhang, Jie
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20020150898A1el Nucleic Acids and
; FILE REFERENCE: 791CIP2E
; CURRENT APPLICATION NUMBER: US/09/816,828
; CURRENT FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: 09/770,160
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 09/552,929
; PRIOR FILING DATE: 2000-04-18
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: pt_FL-genes Version 2.0
```

```
; SEQ ID NO 9
; LENGTH: 2329
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1612)..(2211)
US-09-816-828-9

Query Match      17.5%; Score 36.4; DB 10; Length 2329;
Best Local Similarity 51.9%; Pred. No. 0.44;
Matches 82; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

QY 51 AAGAAGCTCGGCGCTGAACACCGGCTCCCGATCACCAGCAGCACTCTGACCCGAAGAGGAC 110
Db 1924 AAGGGCCACGCAAGAGGTGGCCGCGCTGACCAACGCCGCTGGCGCACGTGGAGCAGC 1983
QY 111 GTCGTCGCCACCATCGAGTACCTGGTCCGCTGCAGGAGGGCCACACGATGACCGTGC 170
Db 1984 ATGCCCAACCGCGTGGCGCACGTGGAGCAGCATGCCCAACGCGCTGTCCGCGCTGAGCGC 2043
QY 171 CCGGGCGGAGTGCAGGTGCGCGTGGAAACGAGCAGCAT 208
Db 2044 CTGCACGCGCACAAAGCTTCGGGTGGACCCGGTCAACTT 2081

RESULT 13
US-09-815-242-7788
; Sequence 7788, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7788
; LENGTH: 1704
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1704)
US-09-815-242-7788

Query Match      17.4%; Score 36.2; DB 10; Length 1704;
Best Local Similarity 50.3%; Pred. No. 0.48;
Matches 89; Conservative 0; Mismatches 88; Indels 0; Gaps 0;
```





GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 12, 2002, 15:56:36 ; Search time 27.8936 Seconds  
(without alignments)  
2286.857 Million cell updates/sec

Title: US-09-697-123B-17  
Perfect score: 208  
Sequence: 1 tcaaggagaacgctacgac.....ccggtggaaaccgacgacat 208

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents\_NA.\*  
1: /cgn2\_6/ptodata/1/ina/5A.COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B.COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A.COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B.COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PCTUS.COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	160	76.9	970	1	US-08-250-030-1
2	160	76.9	970	5	PCT-US95-08790-1
3	144	69.2	3447	2	US-08-313-185-57
4	144	69.2	3447	3	US-09-082-614A-57
5	38.6	18.6	1575	1	US-07-988-2608-2
6	38.2	18.4	2122	4	US-09-029-603-1
7	38.2	18.4	44377	2	US-08-804-227C-7
8	38.2	18.4	44377	2	US-08-804-198-1
9	38	18.3	717	4	US-09-513-783A-49
10	38	18.3	720	3	US-09-094-359-7
11	38	18.3	720	3	US-09-172-063-13
12	38	18.3	720	4	US-09-316-919-14
13	38	18.3	768	3	US-09-094-359-11
14	38	18.3	972	3	US-09-172-063-27
15	38	18.3	1623	4	US-09-513-783A-33
16	38	18.3	2742	4	US-09-232-468A-1
17	38	18.3	4897	6	5196516-7
18	37.2	17.9	942	3	US-08-732-412-1
19	37	17.8	2929	4	US-09-543-084A-31
20	37	17.8	3193	4	US-09-543-084A-30
21	37	17.8	3222	4	US-09-543-084A-29
22	37	17.8	4050	4	US-09-543-084A-26
23	37	17.8	4093	4	US-09-543-084A-28
24	37	17.8	4101	4	US-09-543-084A-27
25	36.4	17.5	1153	4	US-09-372-448A-5
26	36.2	17.4	1894	4	US-09-329-350-32
27	36.2	17.4	3300	1	US-08-194-290-6
28	36.2	17.4	3300	2	US-08-614-377A-6
29	36.2	17.4	3300	4	US-09-142-648B-6
30	36.2	17.4	4403765	4	US-09-103-840A-2
31	36.2	17.4	4411529	4	US-09-103-840A-1
32	36	17.3	765	3	US-08-718-904-79
33	35.8	17.2	13842	4	US-09-105-537-30
34	35.8	17.2	36778	4	US-09-105-537-5
35	35.8	17.2	38506	3	US-09-320-878-19
36	35.4	17.0	6085	4	US-09-029-603-4
37	35.2	16.9	1095	2	US-08-851-088-3
38	35.2	16.9	5535	1	US-08-089-755A-1
39	35.2	16.9	5535	1	US-08-089-755A-4
40	35.2	16.9	5535	1	US-08-421-754-1
41	35.2	16.9	5535	2	US-08-421-754-4
42	35.2	16.9	5535	2	US-08-421-791-1
43	35.2	16.9	5535	2	US-08-421-791-4
44	35	16.8	5011	1	US-08-141-893-1
45	35	16.8	5011	1	US-08-463-092B-1

## ALIGNMENTS

## RESULT 1

US-08-250-030-1  
; Sequence 1, Application US/08250030  
; Patent No. 5643723  
; GENERAL INFORMATION:  
; APPLICANT: Persing, David H.  
; TITLE OF INVENTION: Detection of a Genetic Locus Encoding  
; TITLE OF INVENTION: Resistance to Rifampin in Mycobacterial Cultures and in  
; TITLE OF INVENTION: Clinical Specimens  
; NUMBER OF SEQUENCES: 15  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Schwegman, Lundberg & Woessner  
; STREET: 3500 IDS Center  
; CITY: Minneapolis  
; STATE: MN  
; COUNTRY: USA  
; ZIP: 55402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/250,030  
; FILING DATE: 26-MAY-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Muetting, Ann M.  
; REGISTRATION NUMBER: 33,977  
; REFERENCE/DOCKET NUMBER: 150.105US1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 612-339-0331  
; TELEFAX: 612-339-3061  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 970 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA  
US-08-250-030-1

Query Match 76.9%; Score 160; DB 1; Length 970;  
Best Local Similarity 85.6%; Pred. No. 5e-30;  
Matches 178; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

Qy 1 TCAAGGAGAGCGCTACGACTCGCGCGGTGGCGCTACAGGTCAACAAGAGCTCG 60  
|||||  
Db 26 TCAAGGAGAGCGCTACGACTCGCGCGGTGGCGCTACAGGTCAACAAGAGCTCG 85



; Sequence 57, Application US/09082614A  
; Patent No. 6124098  
; GENERAL INFORMATION:  
; APPLICANT: Heym, Beate  
; APPLICANT: Cole, Stewart  
; APPLICANT: Young, Douglas  
; APPLICANT: Zhang, Ying  
; APPLICANT: Honore, Nadine  
; APPLICANT: Telenti, Annali  
; APPLICANT: Bodmer, Thomas  
; TITLE OF INVENTION: Rapid Detection of Antibiotic Resistance  
; TITLE OF INVENTION: In Mycobacterium Tuberculosis  
; NUMBER OF SEQUENCES: 66  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &  
; ADDRESSEE: Dunner  
; STREET: 1300 I Street, N.W.  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20005-3315  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/082,614A  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/313,185  
; FILING DATE: 12-OCT-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Meyers, Kenneth J.  
; REGISTRATION NUMBER: 25,146  
; REFERENCE/DOCKET NUMBER: 02356.0068-00000  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 408-4000  
; TELEFAX: (202) 408-4400  
; INFORMATION FOR SEQ ID NO: 57:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3447 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
US-09-082-614A-57

Query Match 69.2%; Score 144; DB 3; Length 3447;  
Best Local Similarity 80.8%; Pred. No. 3.6e-26;  
Matches 168; Conservative 0; Mismatches 40; Indels 0; Gaps 0;  
QY 1 TCAAGGAGAACCGGTACGACCTCGCGGGTGGCCCTACAAGGTCACAAAGAGCTCG 60  
Db 809 TCAAGGAGAACCGGTACGACCTCGCGGGTGGCCCTACAAGGTCACAAAGAGCTCG 60  
QY 61 GCCTGAACACCGCGTCCCGCATCAGACGACCACTCTGACCGAAGAGGAGCTCGTCCCA 120  
Db 869 GTTTGCACGCGGGTGAAGTTCATCAGCTGCTCCAGCTGACCGAAGAGGATGTGTCGCCA 928  
QY 121 CCATCGAGTACCTGGTCCGCTGCACGAGGGCCACACGATCACCGTCCCGGGGGAG 180  
Db 929 CCATAGAGTACCTGGTCCGCTGCATGAGGTGAGTGCAGCAATGACTGTCCCGAGGTGGG 988  
QY 181 TCGAGGTGCGCGTGAACACGACAT 208  
Db 989 TAGAAGTGCCAGTGAAACTGACGATAT 1016

RESULT 5  
US-07-988-260B-2  
; Sequence 2, Application US/07988260B

; Patent No. 5482843  
; GENERAL INFORMATION:  
; APPLICANT: BRZEZINSKI, RYSZARD  
; TITLE OF INVENTION: ENZYME OF USE IN CHITOSAN  
; TITLE OF INVENTION: HYDROLYSIS  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: CUSHMAN, DARBY & CUSHMAN  
; STREET: 1100 New York Avenue, N.W., 9th Floor  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005-3918  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 inch, 720 kb diskette  
; COMPUTER: IBM PS/2, Model 30  
; OPERATING SYSTEM: PC-DOS 3.30  
; SOFTWARE: WordPerfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/988,260B  
; FILING DATE: 19921214  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Hymo, Lawrence A.  
; REGISTRATION NUMBER: 19,057  
; REFERENCE/DOCKET NUMBER: LAH/3122/98214/MJW  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 861-3000  
; TELEFAX: (202) 822-0944  
; TELEX: 6714627 CUSH  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1575 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA to mRNA  
; ORIGINAL SOURCE: Streptomyces N174  
US-07-988-260B-2  
Query Match 18.6%; Score 38.6; DB 1; Length 1575;  
Best Local Similarity 51.4%; Pred. No. 0.41;  
Matches 89; Conservative 0; Mismatches 84; Indels 0; Gaps 0;  
QY 14 CTACGACCTCGCGGGTGGCGGTACAGGTCAACAGAGCTCGGCTGAACACCGC 73  
Db 776 CGACGCCATCGTGATGCACGCGCCCGGCAACGACCGACGCTTCGGTGGCATCCGAA 835  
QY 74 GTCCCGCATCAGCAGCACCCTCTGACCGAAGAGGAGCTGCTGCCACCATCGACTCT 133  
Db 836 GACCGCATGAGAGGCGGACGACCCCGCCCGGCGGCGAGAGACCATCTCA 895  
QY 134 GTTCGCGCTGCACGAGGGCCACACGATGACCTCCCGGGCGGAGTCGAGG 186  
Db 896 CGCCTTCTCTGGACGCGCGGCGCCCATGTCTCACCAGGCGCGCGCAGGAG 948  
RESULT 6  
US-09-029-603-1  
; Sequence 1, Application US/09029603  
; Patent No. 6210935  
; GENERAL INFORMATION:  
; APPLICANT: Schupp, Thomas  
; APPLICANT: Engel, Natalie  
; APPLICANT: Bietenhader, Jurg  
; APPLICANT: Toupet, Christine  
; APPLICANT: Pospiech, Andreas  
; TITLE OF INVENTION: Staurosporin Biosynthesis Gene Clusters  
; FILE REFERENCE: 4-20555/A/PCT  
; CURRENT APPLICATION NUMBER: US/09/029,603  
; CURRENT FILING DATE: 1998-03-20  
; EARLIER APPLICATION NUMBER: PCT/EP96/03643  
; EARLIER FILING DATE: 1996-08-19

; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: Patent Ver. 2.0  
; SEQ ID NO 1  
; TYPE: DNA  
; LENGTH: 2122  
; ORGANISM: Streptomyces longisporoflavus  
; FEATURE:  
; NAME/KEY: misc\_RNA  
; LOCATION: (1)..(2122)  
; OTHER INFORMATION: product = 2.1 kb region  
US-09-029-603-1

Query Match 18.4%; Score 38.2; DB 4; Length 2122;  
Best Local Similarity 50.8%; Pred. No. 0.52;  
Matches 91; Conservative 0; Mismatches 88; Indels 0; Gaps 0;  
QY 8 GAAGCGCTACGACCTCGCGCGGTGGCGGTACAAAGCTCAACAGAGCTCGCGCTGAA 67  
Db 1470 GCACCGCTCGGGCTCGCGCTCGCGGAGATCTGACGTCACCGAGACGACGCTGCG 1529  
QY 68 CACCGCGTCCCGATCAGAGACGACACTCTGACCGAAGAGGACGTCGTCGCCACCATCGA 127  
Db 1530 CCTCGCCGACGAGATCGGAGCTCGCGCGTCTGAGGAGCGCCCGTGGCCATGGACGA 1589  
QY 128 GTACCTGGTCCGCTGCGAGGAGGCCACACCATGACCGTCCCGGGCGGAGTCGAGG 186  
Db 1590 GGGCAACTTCGCGCTTCGGCGACGACTCTTCAAGCCGTCGACCTGGCGGCGTCGACG 1648

RESULT 7  
US-08-804-227C-7  
; Sequence 7, Application US/08804227C  
; Patent No. 5876991  
; GENERAL INFORMATION:  
; APPLICANT: DeHoff, Bradley S.  
; APPLICANT: Kuhstoss, Stuart A.  
; APPLICANT: Rostock, Paul R., Jr.  
; APPLICANT: Sutton, Kimberly L.  
; TITLE OF INVENTION: POLYKETIDE SYNTHASE GENES  
; NUMBER OF SEQUENCES: 15  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: THOMAS G. PLANT 1501  
; STREET: LILLY CORPORATE CENTER  
; CITY: INDIANAPOLIS  
; STATE: IN  
; COUNTRY: USA  
; ZIP: 46285  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: ASCII(DOS) Text only  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/804,227C  
; FILING DATE: February 21, 1997  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Plant, Thomas, G.  
; REGISTRATION NUMBER: 35,784  
; REFERENCE/DOCKET NUMBER: X-8231  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 317-276-2459  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 44377 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 350..14002  
; FEATURE:

; NAME/KEY: CDS  
; LOCATION: 14046..20036  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 20110..31284  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 31329..36071  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 36155..41830  
US-08-804-227C-7

Query Match 18.4%; Score 38.2; DB 2; Length 44377;  
Best Local Similarity 49.3%; Pred. No. 0.69;  
Matches 100; Conservative 0; Mismatches 103; Indels 0; Gaps 0;  
QY 5 GGAGAAGCGCTACGACCTCGCGCGGTGGCGGTGCGCTACAAGGTCAACAAGAAGCTCGGCCT 64  
Db 20997 GGTGCGGGGCGCGGTCAACCCAGGACGTCGCCACGACGGCTCACCGCACCACGCGG 21056  
QY 65 GAACACCGCTCCCGATCAGAGACGACCACTGTGACCGAAGAGAGAGCTGTGCGCCACCAT 124  
Db 21057 GCGCGCCGACGACGCGGTGATCGGTGACGCGTGGCGGCGTACGCGCGGCTGACGCGCGCGCA 21116  
QY 125 CGAGTACCTGTCGCGCTGTCACGAGGCGCACACGACGATGACGTCGCGGGGAGTCTGA 184  
Db 21117 CGTGAGCGGTCGAGGCGCACGGCACCGCTCGGGGACCCGATCGAGGCGG 21176  
QY 185 GGTGCGCGGTGGAACCGACGACA 207  
Db 21177 CGCGCTGATGGCCACCTACGGCA 21199

RESULT 8  
US-08-804-198-1  
; Sequence 1, Application US/08804198  
; Patent No. 5945320  
; GENERAL INFORMATION:  
; APPLICANT: Burgett, Stanley G.  
; APPLICANT: Kuhstoss, Stuart A.  
; APPLICANT: Rao, Nagaraja R.  
; APPLICANT: Richardson, Mark A.  
; APPLICANT: Rostock, Paul R., Jr.  
; TITLE OF INVENTION: PLATENOLIDE SYNTHASE GENE  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: PAUL R. CANTRELL 1138  
; STREET: LILLY CORPORATE CENTER  
; CITY: INDIANAPOLIS  
; STATE: IN  
; COUNTRY: USA  
; ZIP: 46285  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: Macintosh  
; OPERATING SYSTEM: Macintosh 7.0  
; SOFTWARE: Microsoft Word 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/804,198  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CANTRELL, PAUL R.  
; REGISTRATION NUMBER: 36,470  
; REFERENCE/DOCKET NUMBER: P9113  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 317-276-3885  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 44377 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single

```
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 350..14002
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 14046..20036
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 20110..31284
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 31329..36071
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 36155..41830
; US-08-804-198-1

Query Match      18.4%; Score 38.2; DB 2; Length 44377;
Best Local Similarity 49.3%; Pred. No. 0.69;
Matches 100; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

QY 5 GGAGAAGCGCTACGACCTCGCGCGGTGGCGGTACAAAGTCAACAAAGAGTCTCGGCGCT 64
Db 20997 GGTGCGGGGCGAGCGGGTCAACACGAGGAGGTGCCAGCAACGGCTCACCGCACCCAGCGG 21056

QY 65 GAACACCCCGTCCCGATCAGCAGCACCTCTGACCGAAGAGGACGCTGTCGCCACCAT 124
Db 21057 GCCCGCCAGCAGGGGTGATCCGTGAGCGCTGGCGGACGCGGGGTGAGCGCCGCGA 21116

QY 125 CGAGTACCTGTGCGCCTGACGAGGCCACACACGATGACCTCCCGCGCGGAGTCTGA 184
Db 21117 CGTGACGCGTTCGAGGCGCACGCGCACCCGCGTCCGCGACCGCGATCGAGGCGG 21176

QY 185 GGTGCGGTGGAACCCGACGACA 207
Db 21177 CGCGCTGATGCCACCTACGGCA 21199

RESULT 9
US-09-513-783A-49
; Sequence 49, Application US/09513783A
; Patent No. 6416959
; GENERAL INFORMATION:
; APPLICANT: Giuliano, Kenneth A.
; APPLICANT: Kapur, Ravi
; TITLE OF INVENTION: A System for Cell Based Screening
; FILE REFERENCE: 97-022-L1
; CURRENT APPLICATION NUMBER: US/09/513,783A
; CURRENT FILING DATE: 2000-02-25
; NUMBER OF SEQ ID NOS: 180
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 49
; LENGTH: 717
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(717)
; OTHER INFORMATION: Description of Artificial Sequence: ECFP
; US-09-513-783A-49

Query Match      18.3%; Score 38; DB 4; Length 717;
Best Local Similarity 49.0%; Pred. No. 0.53;
Matches 101; Conservative 0; Mismatches 105; Indels 0; Gaps 0;

QY 1 TCAAGGAGAAGCGCTACGACCTCGCGGGTGGCGGTACAAAGTCAACAAGAGCTCG 60
Db 392 TCAAGGAGGACGCAACATCTCTGGGCGACAAAGCTGGAGTACAACTACATCAGCCAAAG 451

QY 61 GCTGNACACCGGTCCCGGATCAGCAGCACCACTCTGACCGAAGAGGAGCTGTCGCGCA 120
Db 452 TCTATATACACCGCGACAGCAGAGAGCGCATCAAGGCCCACTTCAAGATCCCGCCACA 511

QY 121 CCATCGAGTACCTGGTCCGCTGACGAGGCGCCACACGATGACCTCCCGGGCGGAG 180
Db 512 ACATCGAGGACGCGAGCTGTCAGCTCGCGGACCACTACGAGCAACACCCCATCGCGG 571

QY 181 TCGAGGTGCGCGTGGAAACCGAGAC 206
Db 572 ACGGCCCGCTGCTGCTGCCGCGACAAC 597

RESULT 11
US-09-172-063-13
; Sequence 13, Application US/09172063
; Patent No. 6150176
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger Y.
; APPLICANT: Miyawaki, Atsushi
; APPLICANT: Llopis, Juan
; APPLICANT: Wachter, Rebekka M.
; APPLICANT: Remington, S. James
; TITLE OF INVENTION: FLUORESCENT PROTEIN SENSORS FOR
; FILE REFERENCE: 07257/071001
; CURRENT APPLICATION NUMBER: US/09/172,063
; CURRENT FILING DATE: 1998-10-13
; EARLIER APPLICATION NUMBER: 09/094,359
```

```
Db 452 TCTATATCAGCGCGGCAAGCAAGAAACGGCATCAAGGCCAACTTCAAGATCCGCCACA 511
QY 121 CCATCGAGTACCTGGTCCGCTGCACGAGGCCACACGATGACCGTCCCGGGCGGAG 180
Db 512 ACATCGAGGACGCGAGCTGCTGCCGCCACTACGAGCAACACCCCATCGCGG 571
QY 181 TCGAGGTGCGCGTGGAAACCGAGAC 206
Db 572 ACGGCCCGCTGCTGCTGCCGCGACAAC 597

RESULT 10
US-09-094-359-7
; Sequence 7, Application US/09094359
; Patent No. 6140132
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger Y.
; APPLICANT: Miyawaki, Atsushi
; APPLICANT: Llopis, Juan
; TITLE OF INVENTION: FLUORESCENT PROTEIN SENSORS FOR
; FILE REFERENCE: 07257/067001
; CURRENT APPLICATION NUMBER: US/09/094,359
; CURRENT FILING DATE: 1998-06-09
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 720
; TYPE: DNA
; ORGANISM: Aequorea victoria
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(717)
; OTHER INFORMATION: ECFP
; US-09-094-359-7

Query Match      18.3%; Score 38; DB 3; Length 720;
Best Local Similarity 49.0%; Pred. No. 0.53;
Matches 101; Conservative 0; Mismatches 105; Indels 0; Gaps 0;

QY 1 TCAAGGAGAAGCGCTACGACCTCGCGCGGTGGCGGTACAAAGTCAACAAGAGCTCG 60
Db 392 TCAAGGAGGACGCAACATCTCTGGGCGACAAAGCTGGAGTACAACTACATCAGCCAAAG 451

QY 61 GCTGNACACCGGTCCCGGATCAGCAGCACCACTCTGACCGAAGAGGAGCTGTCGCGCA 120
Db 452 TCTATATCAGCGCGACAGCAGAGAGCGCATCAAGGCCCACTTCAAGATCCCGCCACA 511

QY 121 CCATCGAGTACCTGGTCCGCTGACGAGGCGCCACACGATGACCTCCCGGGCGGAG 180
Db 512 ACATCGAGGACGCGAGCTGTCAGCTCGCGGACCACTACGAGCAACACCCCATCGCGG 571

QY 181 TCGAGGTGCGCGTGGAAACCGAGAC 206
Db 572 ACGGCCCGCTGCTGCTGCCGCGACAAC 597

RESULT 11
US-09-172-063-13
; Sequence 13, Application US/09172063
; Patent No. 6150176
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger Y.
; APPLICANT: Miyawaki, Atsushi
; APPLICANT: Llopis, Juan
; APPLICANT: Wachter, Rebekka M.
; APPLICANT: Remington, S. James
; TITLE OF INVENTION: FLUORESCENT PROTEIN SENSORS FOR
; FILE REFERENCE: 07257/071001
; CURRENT APPLICATION NUMBER: US/09/172,063
; CURRENT FILING DATE: 1998-10-13
; EARLIER APPLICATION NUMBER: 09/094,359
```



; OTHER INFORMATION: GT-ECFP  
US-09-172-063-27

Query Match	18.38;	Score 38;	DB 3;	Length 972;
Best Local Similarity	49.08;	Pred. No. 0.54;		
Matches 101;	Conservative 0;	Mismatches 105;	Indels 0;	Gaps 0;
Qy 1	TCAAGGAGAGCGCTACGACCTCGCGGGTGGCCGCTACAAGGTCAACAAGAGCTCG	60		
Db 644	TCAAGGAGGAGCGCAACATCCTGGGCAACGTGGAGTACAACACTACGCGCAACG	703		
Qy 61	GCCTGAACACCGGTGCCCGATACGAGGAGCACCTCTGACCGAAGAGGACGCTGTCGCCA	120		
Db 704	TCTATTATCACCGCCGACAAAGCAGAGAAGAACGGCATCAAGGGCCCATCTCAAGATCCGCCACA	763		
Qy 121	CCATCGAGTACCTGGTCCGCTGCACGAGGGCCACACACGATGACCGTCCCGGGCGGAG	180		
Db 764	ACATCGAGGAGCGACGCTGCAGCTGCCGACCACTACCAGCAAAACCCCCCATCGCGG	823		
Qy 181	TCGAGGTCCGGTGGAAACCGACGAC	206		
Db 824	ACGGCCCGGTGCTGCTGCCGACAAAC	849		

RESULT 15

```

US-09-513-783A-33
; Sequence 33, Application US/09513783A
; Patent No. 6416959
; GENERAL INFORMATION:
; APPLICANT: Giuliano, Kenneth A.
; APPLICANT: Kapur, Ravi
; TITLE OF INVENTION: A System for Cell Based Screening
; FILE REFERENCE: 97-022-11
; CURRENT APPLICATION NUMBER: US/09/513,783A
; CURRENT FILING DATE: 2000-02-25
; NUMBER OF SEQ ID NOS: 180
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 33
; LENGTH: 1623
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1623)
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: VFP-NLS-CP3-multiple DEVD-CFP-Annexin II construct
US-09-513-783A-33

```

Query Match	18.3%	Score 38	DB 4	Length 1623
Best Local Similarity	49.0%	Pred. No. 0.57		
Matches 101	Conservative	0	Mismatches 105	Indels 0
Gaps	0			
QY 1	TCAAGGAAAGCGCTACGACCTCGCGGGGTGGCCGCTACAAGGTCAACAAGAAGCTCG	60		
Db 1223	TCAAGGAGGACGGCAACATCTCTGGGCACAAGCTGGAGTACAATCATACGCCACAACG	1282		
QY 61	GCGTGAACACCGCGTCCCGGATACGACGACCGACCTCTGACCGAAGAGGAGCTGCTGCACA	120		
Db 1283	TCATATATCACCCGCACAAAGCAGAGAAGACGGCATCAAGGCCAACTTCAAGATCCGCCACA	1342		
QY 121	CCATCGAGTACCTGTGTCGGCTGCACAGGCGCCACACCCAGCATGCCGTCGCGGGCGGAG	180		
Db 1343	ACATCGAGCAGCGGAGCGTGCAGTCTGCCGACCATACCAGGAAACACCCCATTCGGCG	1402		
QY 181	TCGAGGTCCGGTGAACACGACAC	206		
Db 1403	ACGGGCCGCTGTCTGTGCCGACAAAC	1428		

Search completed: November 12, 2002, 20:15:24  
Job time : 81.8936 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 12, 2002, 15:44:35 ; Search time 136.222 Seconds  
(without alignments)  
3438.621 Million cell updates/sec

Title: US-09-697-123B-17

Perfect score: 208

Sequence: 1 tcaaggagaagcgctacgac.....ccggtggaaccgacgacat 208

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_101002.\*

1: /SID52/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*  
2: /SID52/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*  
3: /SID52/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*  
4: /SID52/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*  
5: /SID52/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*  
6: /SID52/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*  
7: /SID52/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*  
8: /SID52/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*  
9: /SID52/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*  
10: /SID52/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*  
11: /SID52/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*  
12: /SID52/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*  
13: /SID52/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*  
14: /SID52/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*  
15: /SID52/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*  
16: /SID52/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*  
17: /SID52/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*  
18: /SID52/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*  
19: /SID52/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*  
20: /SID52/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*  
21: /SID52/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*  
22: /SID52/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*  
23: /SID52/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*  
24: /SID52/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	208	100.0	208	AAS05217	Mycobacterium cela
2	174.4	83.8	208	AAS05208	Mycobacterium kans
3	168	80.8	208	AAS05215	Mycobacterium aviu
4	168	80.8	208	AAS05219	Mycobacterium fort
5	166.4	80.0	208	AAS05201	Mycobacterium gord
6	164.8	79.2	208	AAS05211	Mycobacterium mari
7	164.8	79.2	208	AAS05224	Mycobacterium xeno
8	161.6	77.7	208	AAS05203	Mycobacterium gord
9	161.6	77.7	208	AAS05210	Mycobacterium ulce

10	160	76.9	208	22	AAS05205	Mycobacterium tube
11	160	76.9	208	22	AAS05206	Mycobacterium terr
12	160	76.9	208	22	AAS05216	Mycobacterium bovi
13	160	76.9	970	17	AAT09676	Mycobacterium tube
14	160	76.9	3519	22	AAS051976	Mycobacterium tube
15	160	76.9	3534	22	AAH02079	Mycobacterium tube
16	160	76.9	3853	21	AAA74651	Mycobacterium tube
17	160	76.9	3853	21	AAA89994	M. tuberculosis rp
18	158.4	76.2	208	22	AAS05222	Mycobacterium afri
19	156.8	75.4	208	22	AAS05202	Mycobacterium gord
20	156.8	75.4	208	22	AAS05223	Mycobacterium haem
21	148	71.2	207	22	AAS05212	Mycobacterium szul
22	146.4	70.4	207	22	AAS05204	Mycobacterium gord
23	145.6	70.0	205	22	AAS05220	Mycobacterium intr
24	144	69.2	3447	14	AAO51532	M.leprae rpoB gene
25	142.2	68.4	211	22	AAS05218	Mycobacterium flav
26	140.8	67.7	214	22	AAS05214	Mycobacterium malm
27	140.8	67.7	214	22	AAS05221	Mycobacterium absc
28	138.2	66.4	223	22	AAS05213	Mycobacterium gast
29	134.4	64.6	214	22	AAS05207	Mycobacterium chel
30	128.6	61.8	223	22	AAS05209	Mycobacterium scro
31	81.6	39.2	3495	22	AAH65512	C glutamicum codin
32	81.6	39.2	34980	22	AAH68525	C glutamicum codin
33	69.2	33.3	27426	23	AAS59541	Propionibacterium
34	42.2	20.3	2918	21	AA243922	S. tendae nikkomyc
35	42.2	20.3	2919	21	AA244486	Streptomyces tende
36	41.6	20.0	2826	24	ABQ90502	M. capsulatus gene
37	41.6	20.0	49999	24	ABQ90984	M. capsulatus gene
38	41.4	19.9	1041	21	AA251708	Burkholderia cepac
39	41.4	19.9	1041	21	AA251709	Burkholderia cepac
40	41.4	19.9	1680	24	ABL39997	Synthetic RT polyn
41	41.4	19.9	1978	24	ABL39991	Synthetic protease
42	41.4	19.9	3009	24	ABL40024	Synthetic Pol poly
43	41.4	19.9	3015	24	ABL39983	Synthetic Pol poly
44	41.2	19.8	2205	24	ABL50199	Human RapiA, c-Raf
45	41.2	19.8	2205	24	ABL50231	Human RapiA, c-Raf

#### ALIGNMENTS

RESULT 1  
AAS05217  
ID AAS05217 standard; DNA; 208 BP.  
AC AAS05217;  
XX  
XX  
DT 07-SEP-2001 (first entry)  
XX  
XX  
DE Mycobacterium celatum rpoB gene fragment.  
XX

XX Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;  
XX PCR-restriction fragment length polymorphism analysis; ds.  
XX

OS Mycobacterium celatum.

XX  
XX WO200131061-A1.  
XX

PD 03-MAY-2001.  
XX

XX 27-OCT-2000; 2000WO-KR01223.  
XX

XX 27-OCT-1999; 99KR-0046795.  
XX  
XX (ERUM-) ERUME BIOTECH CO LTD.

XX  
XX Lee H, Park YK, Bai G, Kim S, Kim Y, Park HJ;  
XX  
XX WPI; 2001-300520/31.  
XX

XX New DNA fragments from the rpoB gene of mycobacteria, useful for  
XX diagnosis and identification of many mycobacterial species by  
XX restriction fragment length polymorphism -  
XX

PS Claim 1; Page 45; 50pp; English.  
 XX The present sequence for *Mycobacterium celatum* rpoB gene  
 CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from  
 CC various *Mycobacterium* species. These rpoB gene fragments can be used  
 CC in the diagnosis and identification of *Mycobacterium* species using a  
 CC novel PCR-restriction fragment length polymorphism analysis (PRA)  
 CC method. The method comprises obtaining a restriction fragment length  
 CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,  
 CC amplifying and digesting the DNA fragment from the microorganism to  
 CC be identified and comparing the RFLP patterns from the known rpoB gene  
 CC fragments with the unidentified fragment. The rpoB gene fragments  
 CC are useful to identify a wide range of *Mycobacterium* species, e.g. for  
 CC diagnosis or to obtain epidemiological and pathogenesis information for  
 CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected  
 CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene  
 CC fragments is rapid, precise, simple and cost effective (only 1 PCR  
 CC required), and can differentiate between many species in a single  
 CC experiment, including those difficult to distinguish by usual biochemical  
 CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for  
 CC detecting specific *Mycobacterium* species.  
 XX  
 SQ Sequence 208 BP; 48 A; 71 C; 64 G; 25 T; 0 other;  
 Query Match 100.0%; Score 208; DB 22; Length 208;  
 Best Local Similarity 100.0%; Pred. No. 5.4e-38;  
 Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 TCAAGGAGAGCGGTACGACCTCGCGGGTGGCGCTACAAGGTCAACAAGAGCTCG 60  
 DB 1 TCAAGGAGAGCGGTACGACCTCGCGGGTGGCGCTACAAGGTCAACAAGAGCTCG 60  
 QY 61 GCCTCAACACCGTCCCGGATCAGCAGCACCTCTGACCGAAGAGACGTCTCGCCA 120  
 DB 61 GCCTGACACCGTCCCGGATCAGCAGCACCTCTGACCGAAGAGACGTCTCGCCA 120  
 QY 121 CCATCGAGTACCTGTCGCGCTGACGAGGGCCACACCATGACCGTCCCGGGCGGAG 180  
 DB 121 CCATCGAGTACCTGTCGCGCTGACGAGGGCCACACCATGACCGTCCCGGGCGGAG 180  
 QY 181 TCGAGGTCCCGTGGAAACCGACGACAT 208  
 DB 181 TCGAGGTCCCGTGGAAACCGACGACAT 208  
 RESULT 2  
 AAS05208  
 ID AAS05208 standard; DNA; 208 BP.  
 XX  
 AC AAS05208;  
 XX  
 DT 07-SEP-2001 (first entry)  
 XX  
 DE *Mycobacterium kansasii* rpoB gene fragment.  
 XX  
 KW Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;  
 KW PCR-restriction fragment length polymorphism analysis; ds.  
 XX  
 OS *Mycobacterium kansasii*.  
 XX  
 PN WO200131061-A1.  
 XX  
 XX 03-MAY-2001.  
 XX  
 PF 27-OCT-2000; 2000WO-KR01223.  
 XX  
 PR 27-OCT-1999; 99KR-0046795.  
 XX  
 PA (ERUM-) ERUME BIOTECH CO LTD.  
 XX  
 PI Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;

XX WPI; 2001-300520/31.  
 XX  
 PT New DNA fragments from the rpoB gene of mycobacteria, useful for  
 PT diagnosis and identification of many mycobacterial species by  
 PT restriction fragment length polymorphism -  
 XX Claim 1; Page 42; 50pp; English.  
 XX  
 CC The present sequence for *Mycobacterium kansasii* rpoB gene  
 CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from  
 CC various *Mycobacterium* species. These rpoB gene fragments can be used  
 CC in the diagnosis and identification of *Mycobacterium* species using a  
 CC novel PCR-restriction fragment length polymorphism analysis (PRA)  
 CC method. The method comprises obtaining a restriction fragment length  
 CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,  
 CC amplifying and digesting the DNA fragment from the microorganism to  
 CC be identified and comparing the RFLP patterns from the known rpoB gene  
 CC fragments with the unidentified fragment. The rpoB gene fragments  
 CC are useful to identify a wide range of *Mycobacterium* species, e.g. for  
 CC diagnosis or to obtain epidemiological and pathogenesis information for  
 CC selection of appropriate therapies, including *M. tuberculosis*, *M. leprae*  
 CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected  
 CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene  
 CC fragments is rapid, precise, simple and cost effective (only 1 PCR  
 CC required), and can differentiate between many species in a single  
 CC experiment, including those difficult to distinguish by usual biochemical  
 CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for  
 CC detecting specific *Mycobacterium* species.  
 XX  
 SQ Sequence 208 BP; 51 A; 65 C; 65 G; 27 T; 0 other;  
 Query Match 83.8%; Score 174.4; DB 22; Length 208;  
 Best Local Similarity 89.9%; Pred. No. 1.9e-30;  
 Matches 187; Conservative 0; Mismatches 21; Indels 0; Gaps 0;  
 QY 1 TCAAGGAGAGCGGTACGACCTCGCGGGTGGCGCTACAAGGTCAACAAGAGCTCG 60  
 DB 1 TCAAGGAGAGCGGTACGACCTCGCGGGTGGCGCTACAAGGTCAACAAGAGCTCG 60  
 QY 61 GCCTCAACACCGTCCCGGATCAGCAGCACCTCTGACCGAAGAGACGTCTCGCCA 120  
 DB 61 GCCTGACACCGTCCCGGATCAGCAGCACCTCTGACCGAAGAGACGTCTCGCCA 120  
 QY 121 CCATCGAGTACCTGTCGCGCTGACGAGGGCCACACCATGACCGTCCCGGGCGGAG 180  
 DB 121 CCATCGAGTATCTGTCGCGCTGACGAGGGCCACGACGATGACCGTCCCGGGCGGAG 180  
 QY 181 TCGAGGTCCCGTGGAAACCGACGACAT 208  
 DB 181 TCGAGGTCCCGTGGAAACCGACGACAT 208  
 RESULT 3  
 AAS05215  
 ID AAS05215 standard; DNA; 208 BP.  
 XX  
 AC AAS05215;  
 XX  
 DT 07-SEP-2001 (first entry)  
 XX  
 DE *Mycobacterium avium* rpoB gene fragment.  
 XX  
 KW Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;  
 KW PCR-restriction fragment length polymorphism analysis; ds.  
 XX  
 OS *Mycobacterium avium*.  
 XX  
 PN WO200131061-A1.  
 XX  
 XX 03-MAY-2001.  
 XX  
 PF 27-OCT-2000; 2000WO-KR01223.

XX PR 27-OCT-1999; 99KR-0046795.  
XX PA (ERUM-) ERUME BIOTECH CO LTD.  
XX PI Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;  
XX DR WPI; 2001-300520/31.  
XX PT New DNA fragments from the rpoB gene of mycobacteria, useful for  
XX PT diagnosis and identification of many mycobacterial species by  
XX PT restriction fragment length polymorphism -  
XX PS Claim 1; Page 44; 50pp; English.  
XX CC The present sequence for Mycobacterium avium rpoB gene  
XX CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from  
XX CC various Mycobacterial species. These rpoB gene fragments can be used  
XX CC in the diagnosis and identification of Mycobacterium species using a  
XX CC novel PCR-restriction fragment length polymorphism analysis (PRA)  
XX CC method. The method comprises obtaining a restriction fragment length  
XX CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,  
XX CC amplifying and digesting the DNA fragment from the microorganism to  
XX CC be identified and comparing the RFLP patterns from the known rpoB gene  
XX CC fragments with the unidentified fragment. The rpoB gene fragments  
XX CC are useful to identify a wide range of Mycobacterium species, e.g. for  
XX CC diagnosis or to obtain epidemiological and pathogenesis information for  
XX CC selection of appropriate therapies, including M. tuberculosis, M. leprae  
XX CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected  
XX CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene  
XX CC fragments is rapid, precise, simple and cost effective (only 1 PCR  
XX CC required), and can differentiate between many species in a single  
XX CC experiment, including those difficult to distinguish by usual biochemical  
XX CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for  
XX CC detecting specific Mycobacterial species.  
XX SQ Sequence 208 BP; 44 A; 69 C; 69 G; 26 T; 0 other;  
Query Match 80.8%; Score 168; DB 22; Length 208;  
Best Local Similarity 88.0%; Pred. No. 5.1e-29;  
Matches 183; Conservative 0; Mismatches 25; Indels 0; Gaps 0;  
QY 1 TCAAGGAGAGCGCTACGACCTCGCGGGGTGGCGCTACAGGTCAACAGAGCTCG 60  
Db 1 TCAAGGAGAGCGCTACGACCTCGCGGGGTGGCGCTACAGGTCAACAGAGCTCG 60  
QY 61 GCCTGAACACCGCGTCCCGGATCAGCAGCACCTCTGACCGAGAGAGCGTCTCGCCA 120  
Db 61 GCCTGACGCGGTGAGCCGATCACCAGCTCGACGCTGACCGAGGAGAGCGTCTCGCCA 120  
QY 121 CCATCGAGTACCTGGTCCCGCTCGACGAGGGCCACACCGATGACCGTCCCGGGCGGAG 180  
Db 121 CCATCGAGTACCTGGTCCCGCTCGACGAGGGTTCAGCCACGATGACCGTCCCGGGCGGCA 180  
QY 181 TCGAGGTGCGGTGGAAACCGACGACAT 208  
Db 181 TCGAGGTGCGGTGGAGACCGACGACAT 208  
RESULT 4  
AAS05219  
ID AAS05219 standard; DNA; 208 BP.  
XX AC AAS05219;  
XX DT 07-SEP-2001 (first entry)  
XX DE Mycobacterium fortuitum rpoB gene fragment.  
XX KW Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;  
XX KW PCR-restriction fragment length polymorphism analysis; ds.  
XX OS Mycobacterium fortuitum.

XX WO200131061-A1.  
XX PD 03-MAY-2001.  
XX PF 27-OCT-2000; 2000WO-KR01223.  
XX PR 27-OCT-1999; 99KR-0046795.  
XX PA (ERUM-) ERUME BIOTECH CO LTD.  
XX PI Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;  
XX DR WPI; 2001-300520/31.  
XX PT New DNA fragments from the rpoB gene of mycobacteria, useful for  
XX PT diagnosis and identification of many mycobacterial species by  
XX PT restriction fragment length polymorphism -  
XX PS Claim 1; Page 46; 50pp; English.  
XX CC The present sequence for Mycobacterium fortuitum rpoB gene  
XX CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from  
XX CC various Mycobacterial species. These rpoB gene fragments can be used  
XX CC in the diagnosis and identification of Mycobacterium species using a  
XX CC novel PCR-restriction fragment length polymorphism analysis (PRA)  
XX CC method. The method comprises obtaining a restriction fragment length  
XX CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,  
XX CC amplifying and digesting the DNA fragment from the microorganism to  
XX CC be identified and comparing the RFLP patterns from the known rpoB gene  
XX CC fragments with the unidentified fragment. The rpoB gene fragments  
XX CC are useful to identify a wide range of Mycobacterium species, e.g. for  
XX CC diagnosis or to obtain epidemiological and pathogenesis information for  
XX CC selection of appropriate therapies, including M. tuberculosis, M. leprae  
XX CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected  
XX CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene  
XX CC fragments is rapid, precise, simple and cost effective (only 1 PCR  
XX CC required), and can differentiate between many species in a single  
XX CC experiment, including those difficult to distinguish by usual biochemical  
XX CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for  
XX CC detecting specific Mycobacterial species.  
XX SQ Sequence 208 BP; 43 A; 69 C; 69 G; 27 T; 0 other;  
Query Match 80.8%; Score 168; DB 22; Length 208;  
Best Local Similarity 88.0%; Pred. No. 5.1e-29;  
Matches 183; Conservative 0; Mismatches 25; Indels 0; Gaps 0;  
QY 1 TCAAGGAGAGCGCTACGACCTCGCGGGGTGGCGCTACAGGTCAACAGAGCTCG 60  
Db 1 TCAAGGAGAGCGCTACGACCTCGCGGGGTGGCGCTACAGGTCAACAGAGCTCG 60  
QY 61 GCCTGAACACCGCGTCCCGGATCAGCAGCACCTCTGACCGAGAGAGCGTCTCGCCA 120  
Db 61 GCCTGACGCGGTGAGCCGATCACCAGCTCGACGCTGACCGAGGAGAGCGTCTCGCCA 120  
QY 121 CCATCGAGTACCTGGTCCCGCTCGACGAGGGCCACACCGATGACCGTCCCGGGCGGAG 180  
Db 121 CCATCGAGTACCTGGTCCCGCTCGACGAGGGCCACACCGATGACCGTCCCGGGCGGAG 180  
QY 181 TCGAGGTGCGGTGGAAACCGACGACAT 208  
Db 181 TCGAGGTGCGGTGGAGACCGACGACAT 208  
RESULT 5  
AAS05201  
ID AAS05201 standard; DNA; 208 BP.  
XX AC AAS05201;  
XX DT 07-SEP-2001 (first entry)  
XX

DE Mycobacterium gordonae type I rpoB gene fragment.

KW Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;  
KW PCR-restriction fragment length polymorphism analysis; ds.

OS Mycobacterium gordonae type I.

PN WO200131061-A1.

PD 03-MAY-2001.

PF 27-OCT-2000; 2000WO-KR01223.

PR 27-OCT-1999; 99KR-0046795.

PA (ERUM-) ERUME BIOTECH CO LTD.

PI Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;

DR WPI; 2001-300520/31.

XX New DNA fragments from the rpoB gene of mycobacteria, useful for  
PT diagnosis and identification of many mycobacterial species by  
PT restriction fragment length polymorphism

PS Claim 1; Page 40; 50pp; English.

CC The present sequence for Mycobacterium gordonae type I rpoB gene  
CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from  
CC various Mycobacterial species. These rpoB gene fragments can be used  
CC in the diagnosis and identification of Mycobacterium species using a  
CC novel PCR-restriction fragment length polymorphism analysis (PRA)  
CC method. The method comprises obtaining a restriction fragment length  
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,  
CC amplifying and digesting the DNA fragment from the microorganism to  
CC be identified and comparing the RFLP patterns from the known rpoB gene  
CC fragments with the unidentified fragment. The rpoB gene fragments  
CC are useful to identify a wide range of Mycobacterium species, e.g. for  
CC diagnosis or to obtain epidemiological and pathogenesis information for  
CC selection of appropriate therapies, including M. tuberculosis, M. leprae  
CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected  
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene  
CC fragments is rapid, precise, simple and cost effective (only 1 PCR  
CC required), and can differentiate between many species in a single  
CC experiment, including those difficult to distinguish by usual biochemical  
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for  
CC detecting specific Mycobacterial species.

XX Sequence 208 BP; 46 A; 71 C; 65 G; 26 T; 0 other;

Query Match 80.0%; Score 166.4; DB 22; Length 208;  
Best Local Similarity 87.5%; Pred. No. 1.2e-28;  
Matches 182; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 1 TCAAGGAGAGCGCTAGGACCTCGCGCGGTGGCGCTACAGGTCAACAGAGCTCG 60

DB 1 TCAAGGAGAGCGCTAGGACCTCGCGCGGTGGCGCTACAGGTCAACAGAGCTCG 60

QY 61 GCCTGAACACCGCTGCCGATCAGGACGACCTCTGACCGAAGAGAGCTCGCGCA 120

DB 61 GCCTGAACACCGCTGCCGATCAGGACGACCTCTGACCGAAGAGAGCTCGCGCA 120

QY 121 CCATCGAGTACCTGGTCCGCTGCAGAGGGCCACACGATGACCGTCCCGGGCGGAG 180

DB 121 CCATCGAGTACCTGGTCCGCTGCAGAGGGCCACACGATGACCGTCCCGGGCGGAG 180

QY 181 TCGAGGTGCGGTGGAAACCGAGACAT 208

DB 181 CCGAGGTGCGGTGGAGCCGAGACAT 208

RESULT 6  
AAS05211

ID

XX

AC

XX

DT

XX

DE

XX

KW

KW

XX

OS

XX

PN

XX

PD

XX

PF

XX

PR

XX

PA

XX

PI

XX

DR

XX

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

AAS05211 standard; DNA; 208 BP.

AAS05211;

07-SEP-2001 (first entry)

Mycobacterium marinum rpoB gene fragment.

Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;  
PCR-restriction fragment length polymorphism analysis; ds.

Mycobacterium marinum.

WO200131061-A1.

03-MAY-2001.

27-OCT-2000; 2000WO-KR01223.

27-OCT-1999; 99KR-0046795.

(ERUM-) ERUME BIOTECH CO LTD.

Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;

WPI; 2001-300520/31.

New DNA fragments from the rpoB gene of mycobacteria, useful for  
diagnosis and identification of many mycobacterial species by  
restriction fragment length polymorphism

Claim 1; Page 43; 50pp; English.

CC The present sequence for Mycobacterium marinum rpoB gene  
CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from  
CC various Mycobacterial species. These rpoB gene fragments can be used  
CC in the diagnosis and identification of Mycobacterium species using a  
CC novel PCR-restriction fragment length polymorphism analysis (PRA)  
CC method. The method comprises obtaining a restriction fragment length  
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,  
CC amplifying and digesting the DNA fragment from the microorganism to  
CC be identified and comparing the RFLP patterns from the known rpoB gene  
CC fragments with the unidentified fragment. The rpoB gene fragments  
CC are useful to identify a wide range of Mycobacterium species, e.g. for  
CC diagnosis or to obtain epidemiological and pathogenesis information for  
CC selection of appropriate therapies, including M. tuberculosis, M. leprae  
CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected  
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene  
CC fragments is rapid, precise, simple and cost effective (only 1 PCR  
CC required), and can differentiate between many species in a single  
CC experiment, including those difficult to distinguish by usual biochemical  
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for  
CC detecting specific Mycobacterial species.

XX Sequence 208 BP; 45 A; 69 C; 68 G; 26 T; 0 other;

Query Match 79.2%; Score 164.8; DB 22; Length 208;  
Best Local Similarity 87.0%; Pred. No. 2.6e-28;  
Matches 181; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

QY 1 TCAAGGAGAGCGCTAGGACCTCGCGCGGTGGCGCTACAGGTCAACAGAGCTCG 60

DB 1 TCAAGGAGAGCGCTAGGACCTCGCGCGGTGGCGCTACAGGTCAACAGAGCTCG 60

QY 61 GCCTGAACACCGCTGCCGATCAGGACGACCTCTGACCGAAGAGAGCTCGCGCA 120

DB 61 GCCTGAACACCGCTGCCGATCAGGACGACCTCTGACCGAAGAGAGCTCGCGCA 120

QY 121 CCATCGAGTACCTGGTCCGCTGCAGAGGGCCACACGATGACCGTCCCGGGCGGAG 180

DB 121 CCATCGAGTACCTGGTCCGCTGCAGAGGGCCACACGATGACCGTCCCGGGCGGAG 180

QY 181 TCGAGGTGCGGTGGAAACCGAGACAT 208

```

Db 181 TCGAGGTCCCGTGCAGACCGACGACAT 208
|||||
RESULT 7
AAS05224
ID AAS05224 standard; DNA: 208 BP.
XX
XX
AC AAS05224;
XX
XX
DT 07-SEP-2001 (first entry)
XX
DE Mycobacterium xenopi rpoB gene fragment.
XX
XX
KW Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;
XX PCR-restriction fragment length polymorphism analysis; ds.
XX
OS Mycobacterium xenopi.
XX
XX
PN WO200131061-A1.
XX
XX
PD 03-MAY-2001.
XX
XX
PF 27-OCT-2000; 2000WO-KR01223.
XX
XX
PR 27-OCT-1999; 99KR-0046795.
XX
XX
PA (ERUM-) ERUME BIOTECH CO LTD.
XX
XX
PI Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;
XX
XX
DR WPI: 2001-300520/31.
XX
XX
PT New DNA fragments from the rpoB gene of mycobacteria, useful for
PT diagnosis and identification of many mycobacterial species by
PT restriction fragment length polymorphism -
XX
XX
PS Claim 1; Page 47; 50pp; English.
XX
XX
CC The present sequence for Mycobacterium xenopi rpoB gene
CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
CC various Mycobacterial species. These rpoB gene fragments can be used
CC in the diagnosis and identification of Mycobacterium species using a
CC novel PCR-restriction fragment length polymorphism analysis (PRA)
CC method. The method comprises obtaining a restriction fragment length
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
CC amplifying and digesting the DNA fragment from the microorganism to
CC be identified and comparing the RFLP patterns from the known rpoB gene
CC fragments with the unidentified fragment. The rpoB gene fragments
CC are useful to identify a wide range of Mycobacterium species, e.g. for
CC diagnosis or to obtain epidemiological and pathogenesis information for
CC selection of appropriate therapies, including M. tuberculosis, M. leprae
CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene
CC fragments is rapid, precise, simple and cost effective (only 1 PCR
CC required), and can differentiate between many species in a single
CC experiment, including those difficult to distinguish by usual biochemical
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
CC detecting specific Mycobacterial species.
XX
SQ Sequence 208 BP; 51 A; 65 C; 67 G; 25 T; 0 other;

Query Match 79.2%; Score 164.8; DB 22; Length 208;
Best Local Similarity 87.0%; Pred. No. 2.6e-28;
Matches 181; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

QY 1 TCAAGGAGAACGGCTACGACCTCGCGGGTGGCGGCTACAGGTCAACAAGAACGCTCG 60
Db 1 TCAAGGAGAACGGCTACGACCTCGCGGGTGGCGGCTACAGGTCAACAAGAACGCTCG 60
QY 61 GCCTGAACACCGCTCCCGGATCAGGACGACCTCTGACCGAGAGAGCGTGTGCGCCA 120
Db 1 TCAAGGAGAACGGCTACGACCTCGCGGGTGGCGGCTACAGGTCAACAAGAACGCTCG 60
QY 61 GCCTGAACACCGAGTACGACCTCGCGGGTGGCGGCTACAGGTCAACAAGAACGCTCG 120

```

```

QY 121 CCATCGAGTACCTGGTCCGCTGCAGAGGGCCACACGACGATGACCGTCCCGGGGGGAG 180
|||||
Db 121 CCATCGAATACCTGGTGGCTTGCAGAGGGCCACCCACGATGAAGGTCCCGGTGGCG 180
|||||
QY 181 TCGAGGTCCCGGTGGAACCGACGACAT 208
|||||
Db 181 TCGAGGTCCCGGTGGAACCGACGACAT 208
|||||

```

## RESULT 8

```

AAS05203
ID AAS05203 standard; DNA: 208 BP.
XX
XX
AC AAS05203;
XX
XX
DT 07-SEP-2001 (first entry)
XX
DE Mycobacterium gordonae type III rpoB gene fragment.
XX
XX
KW Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;
XX PCR-restriction fragment length polymorphism analysis; ds.
XX
OS Mycobacterium gordonae type III.
XX
XX
PN WO200131061-A1.
XX
XX
PD 03-MAY-2001.
XX
XX
PF 27-OCT-2000; 2000WO-KR01223.
XX
XX
PR 27-OCT-1999; 99KR-0046795.
XX
XX
PA (ERUM-) ERUME BIOTECH CO LTD.
XX
XX
PI Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;
XX
XX
DR WPI: 2001-300520/31.
XX

```

```

PT New DNA fragments from the rpoB gene of mycobacteria, useful for
PT diagnosis and identification of many mycobacterial species by
PT restriction fragment length polymorphism -
XX
XX
PS Claim 1; Page 41; 50pp; English.
XX

```

```

CC The present sequence for Mycobacterium gordonae type III rpoB gene
CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
CC various Mycobacterial species. These rpoB gene fragments can be used
CC in the diagnosis and identification of Mycobacterium species using a
CC novel PCR-restriction fragment length polymorphism analysis (PRA)
CC method. The method comprises obtaining a restriction fragment length
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
CC amplifying and digesting the DNA fragment from the microorganism to
CC be identified and comparing the RFLP patterns from the known rpoB gene
CC fragments with the unidentified fragment. The rpoB gene fragments
CC are useful to identify a wide range of Mycobacterium species, e.g. for
CC diagnosis or to obtain epidemiological and pathogenesis information for
CC selection of appropriate therapies, including M. tuberculosis, M. leprae
CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene
CC fragments is rapid, precise, simple and cost effective (only 1 PCR
CC required), and can differentiate between many species in a single
CC experiment, including those difficult to distinguish by usual biochemical
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
CC detecting specific Mycobacterial species.
XX

```

```

SQ Sequence 208 BP; 46 A; 69 C; 63 G; 30 T; 0 other;

```

```

Query Match 77.7%; Score 161.6; DB 22; Length 208;
Best Local Similarity 86.1%; Pred. No. 1.4e-27;
Matches 179; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY 1 TCAAGGAGAACGGCTACGACCTCGCGGGTGGCGGCTACAGGTCAACAAGAACGCTCG 60

```

Db 1 TCAAGGAGAGCGGTACGACCTGGCCCGCTACAAAGGTCAACAAGAGCTCG 60  
QY 61 GCCTGAACACCGCGTCCCGGATCAGCAGCACCCTCTGACCGAGAGAGCGTGTCCGCCA 120  
Db 61 GCCTGACAGTCCGGGATCCGATCACCAGCTCCACGCTGACCGAAGAAGAGCGTGTCCGCCA 120  
QY 121 CCATCGAGTACCTGTCGCCGCTGCACGAGGCGCCACACACGATCACCGTCCCGGGCGGAG 180  
Db 121 CCATCGAGTACCTGTCGGTCTGCACGAGGCTGACGACAGTACCGTTCGGGCGGCCA 180  
QY 181 TCGAGGTCCCGGTGAACACGACGACAT 208  
Db 181 CCGAGTTCCGGTGGAGACCGACGACAT 208  
RESULT 9  
AAS05210  
ID AAS05210 standard; DNA; 208 BP.  
XX AC AAS05210;  
XX DT 07-SEP-2001 (first entry)  
XX DE Mycobacterium ulcerans rpoB gene fragment.  
XX KW Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;  
XX KW PCR-restriction fragment length polymorphism analysis; ds.  
XX OS Mycobacterium ulcerans.  
XX PN WO200131061-A1.  
XX PD 03-MAY-2001.  
XX PF 27-OCT-2000; 2000WO-KR01223.  
XX PR 27-OCT-1999; 99KR-0046795.  
XX PA (ERUM-) ERUME BIOTECH CO LTD.  
XX PI Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;  
XX WPI; 2001-300520/31.  
XX The present sequence for Mycobacterium ulcerans rpoB gene  
CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from  
CC various Mycobacterial species. These rpoB gene fragments can be used  
CC in the diagnosis and identification of Mycobacterium species using a  
CC novel PCR-restriction fragment length polymorphism analysis (PRA)  
CC method. The method comprises obtaining a restriction fragment length  
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,  
CC amplifying and digesting the DNA fragment from the microorganism to  
CC be identified and comparing the RFLP patterns from the known rpoB gene  
CC fragments with the unidentified fragment. The rpoB gene fragments  
CC are useful to identify a wide range of Mycobacterium species, e.g. for  
CC diagnosis or to obtain epidemiological and pathogenesis information for  
CC selection of appropriate therapies, including M. tuberculosis, M. leprae  
CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected  
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene  
CC fragments is rapid, precise, simple and cost effective (only 1 PCR  
CC required), and can differentiate between many species in a single  
CC experiment, including those difficult to distinguish by usual biochemical  
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for  
CC detecting specific Mycobacterial species.  
XX Sequence 208 BP: 45 A; 68 C; 67 G; 28 T; 0 other;

Query Match 77.7%; Score 161.6; DB 22; Length 208;  
Best Local Similarity 86.1%; Pred. No. 1.4e-27;  
Matches 179; Conservative 0; Mismatches 29; Indels 0; Gaps 0;  
QY 1 TCAAGGAGAGCGGTACGACCTGGCCCGGTACAAAGGTCAACAAGAGCTCG 60  
Db 1 TCAAGGAGAGCGGTACGACCTGGCTCGGTCGGTGGGTACAAAGGTCAACAAGAGCTCG 60  
QY 61 GCCTGAACACCGCGTCCCGGATCAGCAGCACCCTCTGACCGAGAGAGCGTGTCCGCCA 120  
Db 61 GCCTGAACCGCGGCCAGCCCATCACGAGCTCGACGCTGACCGAGGAAGAGCGTGTCCGCCA 120  
QY 121 CCATCGAGTACCTGTCGCCGCTGCACGAGGCGCCACACGATCACCGTCCCGGGCGGAG 180  
Db 121 CCATCGAATACCTGTCGGCTTGCACGAGGCGCAGACCGCGATGACCGTCCGGCGGTG 180  
QY 181 TCGAGGTCCCGGTGAACACGACGACAT 208  
Db 181 TCGAGGTCCCGGTGCGAGACCGAGACAT 208  
RESULT 10  
AAS05205  
ID AAS05205 standard; DNA; 208 BP.  
XX AC AAS05205;  
XX DT 07-SEP-2001 (first entry)  
XX DE Mycobacterium tuberculosis rpoB gene fragment.  
XX KW Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;  
XX KW PCR-restriction fragment length polymorphism analysis; ds.  
XX OS Mycobacterium tuberculosis.  
XX PN WO200131061-A1.  
XX PD 03-MAY-2001.  
XX PF 27-OCT-2000; 2000WO-KR01223.  
XX PR 27-OCT-1999; 99KR-0046795.  
XX PA (ERUM-) ERUME BIOTECH CO LTD.  
XX PI Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;  
XX WPI; 2001-300520/31.  
XX New DNA fragments from the rpoB gene of mycobacteria, useful for  
PT diagnosis and identification of many mycobacterial species by  
PT restriction fragment length polymorphism -  
PS Disclosure; Page 41; 50pp; English.  
XX The present sequence for Mycobacterium tuberculosis rpoB gene  
CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from  
CC various Mycobacterial species. These rpoB gene fragments can be used  
CC in the diagnosis and identification of Mycobacterium species using a  
CC novel PCR-restriction fragment length polymorphism analysis (PRA)  
CC method. The method comprises obtaining a restriction fragment length  
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,  
CC amplifying and digesting the DNA fragment from the microorganism to  
CC be identified and comparing the DNA fragment from the known rpoB gene  
CC fragments with the unidentified fragment. The rpoB gene fragments  
CC are useful to identify a wide range of Mycobacterium species, e.g. for  
CC diagnosis or to obtain epidemiological and pathogenesis information for  
CC selection of appropriate therapies, including M. tuberculosis, M. leprae  
CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected  
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene  
CC fragments is rapid, precise, simple and cost effective (only 1 PCR  
CC required), and can differentiate between many species in a single  
CC experiment, including those difficult to distinguish by usual biochemical  
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for  
CC detecting specific Mycobacterial species.

CC required), and can differentiate between many species in a single  
CC experiment, including those difficult to distinguish by usual biochemical  
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for  
CC detecting specific Mycobacterial species.  
XX  
SQ Sequence 208 BP; 46 A; 62 C; 67 G; 33 T; 0 other;  
Query Match 76.9%; Score 160; DB 22; Length 208;  
Best Local Similarity 85.6%; Pred. No. 3.2e-27;  
Matches 178; Conservative 0; Mismatches 30; Indels 0; Gaps 0;  
QY 1 TCAAGGAGAAGCGCTACGACCTCGCGGGTGGCGCTACAGGTCAACAAGAAGCTCG 60  
DB 1 TCAAGGAGAAGCGCTACGACCTGCGCGCTATAGGTCAACAAGAAGCTCG 60  
QY 61 GCCTGAACACCGCTCCCGNATCAGAGACCACTCTGACCGAAGAGAGCGTCTCGGCCA 120  
DB 61 GCCTGCATGTGCGGAGGCCATCAGCTCGTGCAGCGTGACCGAAGAAGAGCGTCTGGCCA 120  
QY 121 CCATCGAGTACCTGCTCGGCTGCACGAGGCCACACACGATGACCGTCCCGGGCGAG 180  
DB 121 CCATCGAATATCTGCTCGGCTTGACAGAGGTGACAGCACGATGACCGTTCCGGGGCGG 180  
QY 181 TCGAGGTGCCGGTGGAAACCGACGACAT 208  
DB 181 TCGAGGTGCCGGTGGAAACCGACGACAT 208  
RESULT 11  
AAS05206  
ID AAS05206 standard; DNA; 208 BP.  
XX  
AC AAS05206;  
XX  
XX 07-SEP-2001 (first entry)  
XX  
XX Mycobacterium terrae rpoB gene fragment.  
XX  
DE Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;  
KW PCR-restriction fragment length polymorphism analysis; ds.  
XX  
XX Mycobacterium terrae.  
XX  
OS WO200131061-A1.  
XX  
PN 03-MAY-2001.  
XX  
PD 27-OCT-2000; 2000WO-KR01223.  
XX  
PF 27-OCT-1999; 99KR-0046795.  
XX  
PR (ERUM-) ERUME BIOTECH CO LTD.  
XX  
XX Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;  
XX WPI: 2001-300520/31.  
XX  
XX New DNA fragments from the rpoB gene of mycobacteria, useful for  
PT diagnosis and identification of many mycobacterial species by  
PT restriction fragment length polymorphism -  
XX  
XX Claim 1: Page 42; 50pp; English.  
XX  
XX The present sequence for Mycobacterium terrae rpoB gene  
CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from  
CC various Mycobacterial species. These rpoB gene fragments can be used  
CC in the diagnosis and identification of Mycobacterium species using a  
CC novel PCR-restriction fragment length polymorphism analysis (PRA)  
CC method. The method comprises obtaining a restriction fragment length  
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,  
CC amplifying and digesting the DNA fragment from the microorganism to  
CC be identified and comparing the RFLP patterns from the known rpoB gene  
CC fragments with the unidentified fragment. The rpoB gene fragments

CC are useful to identify a wide range of Mycobacterium species, e.g. for  
CC diagnosis or to obtain epidemiological and pathogenesis information for  
CC selection of appropriate therapies, including M. tuberculosis, M. leprae  
CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected  
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene  
CC fragments is rapid, precise, simple and cost effective (only 1 PCR  
CC required), and can differentiate between many species in a single  
CC experiment, including those difficult to distinguish by usual biochemical  
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for  
CC detecting specific Mycobacterial species.  
XX  
SQ Sequence 208 BP; 46 A; 62 C; 67 G; 33 T; 0 other;  
Query Match 76.9%; Score 160; DB 22; Length 208;  
Best Local Similarity 85.6%; Pred. No. 3.2e-27;  
Matches 178; Conservative 0; Mismatches 30; Indels 0; Gaps 0;  
QY 1 TCAAGGAGAAGCGCTACGACCTCGCGGGTGGCGCTACAGGTCAACAAGAAGCTCG 60  
DB 1 TCAAGGAGAAGCGCTACGACCTGCGCGCTATAGGTCAACAAGAAGCTCG 60  
QY 61 GCCTGAACACCGCTCCCGNATCAGAGACCACTCTGACCGAAGAGAGCGTCTCGGCCA 120  
DB 61 GCCTGCATGTGCGGAGGCCATCAGCTCGTGCAGCGTGACCGAAGAAGAGCGTCTGGCCA 120  
QY 121 CCATCGAGTACCTGCTCGGCTGCACGAGGCCACACACGATGACCGTCCCGGGCGAG 180  
DB 121 CCATCGAATATCTGCTCGGCTTGACAGAGGTGACAGCACGATGACCGTTCCGGGGCGG 180  
QY 181 TCGAGGTGCCGGTGGAAACCGACGACAT 208  
DB 181 TCGAGGTGCCGGTGGAAACCGACGACAT 208  
RESULT 12  
AAS05216  
ID AAS05216 standard; DNA; 208 BP.  
XX  
AC AAS05216;  
XX  
XX 07-SEP-2001 (first entry)  
XX  
XX Mycobacterium bovis rpoB gene fragment.  
XX  
DE Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;  
KW PCR-restriction fragment length polymorphism analysis; ds.  
XX  
XX Mycobacterium bovis.  
XX  
OS WO200131061-A1.  
XX  
PN 03-MAY-2001.  
XX  
PD 27-OCT-2000; 2000WO-KR01223.  
XX  
PF 27-OCT-1999; 99KR-0046795.  
XX  
PR (ERUM-) ERUME BIOTECH CO LTD.  
XX  
XX Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;  
XX WPI: 2001-300520/31.  
XX  
XX New DNA fragments from the rpoB gene of mycobacteria, useful for  
PT diagnosis and identification of many mycobacterial species by  
PT restriction fragment length polymorphism -  
XX  
XX Claim 1: Page 45; 50pp; English.  
XX  
XX The present sequence for Mycobacterium bovis rpoB gene  
CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from  
CC various Mycobacterial species. These rpoB gene fragments can be used  
CC in the diagnosis and identification of Mycobacterium species using a  
CC novel PCR-restriction fragment length polymorphism analysis (PRA)  
CC method. The method comprises obtaining a restriction fragment length  
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,  
CC amplifying and digesting the DNA fragment from the microorganism to  
CC be identified and comparing the RFLP patterns from the known rpoB gene  
CC fragments with the unidentified fragment. The rpoB gene fragments

CC novel PCR-restriction fragment length polymorphism analysis (PRA)  
 CC method. The method comprises obtaining a restriction fragment length  
 CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,  
 CC amplifying and digesting the DNA fragment from the microorganism to  
 CC be identified and comparing the RFLP patterns from the known rpoB gene  
 CC fragments with the unidentified fragment. The rpoB gene fragments  
 CC are useful to identify a wide range of Mycobacterium species, e.g. for  
 CC diagnosis or to obtain epidemiological and pathogenesis information for  
 CC selection of appropriate therapies, including M. tuberculosis, M. leprae  
 CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected  
 CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene  
 CC fragments is rapid, precise, simple and cost effective (only 1 PCR  
 CC required), and can differentiate between many species in a single  
 CC experiment, including those difficult to distinguish by usual biochemical  
 CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for  
 CC detecting specific Mycobacterial species.

XX Sequence 208 BP; 46 A; 62 C; 67 G; 33 T; 0 other;

Query Match 76.9%; Score 160; DB 22; Length 208;  
 Best Local Similarity 85.6%; Pred. No. 3.2e-27;  
 Matches 178; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

Qy 1 TCAAGGAGAAGCGGTACGACCTCGCGGGTGGCGGCTACAAGGTCAACAAGAGCTCG 60  
 Db 1 TCAAGGAGAAGCGGTACGACCTCGCGGGTGGCGGCTACAAGGTCAACAAGAGCTCG 60  
 Qy 61 GCCTCAACACCGCTCCCGGATACGAGGACCACTCTGACGAGAGGAGCGTCTCGGCCA 120  
 Db 61 GCCTCATGTGCGGAGCCCATACGCTGCTGACGAGGAGAGGAGCGTCTCGGCCA 120  
 Qy 121 CCATCGAGTACCTGTCGCGCTGACGAGGCGCACACCATGATGACCGTCCCGGGCGGAG 180  
 Db 121 CCATCGAATATCTGTCGCTGTCAGGAGGCTCAGACCATGACCGTCCCGGGCGGCG 180  
 Qy 181 TCGAGGTCCGGTGGAAACCGACGACAT 208  
 Db 181 TCGAGGTCCGGTGGAAACCGACGACAT 208

# RESULT 13

AAT09676 standard; DNA; 970 BP.

AC AAT09676;

DT 15-OCT-1996 (first entry)

XX Mycobacterium tuberculosis rpoB gene DNA sequence.

DE Tuberculosis; disease diagnosis; oligonucleotide; DNA primer; PCR;  
 KW polymerase chain reaction; DNA amplification; rpoB locus; TB; ss.

OS Mycobacterium tuberculosis.

XX Key Location/Qualifiers  
 FH primer\_bind 10..27

FT /tag= a

FT /note= "primer FENLFF"

FT primer\_bind 226..243

FT /tag= b

FT /note= "primer DDIDLH"

FT primer\_bind 226..240

FT /tag= c

FT /note= "primer DDIDLH"

FT primer\_bind 338..364

FT /tag= d

FT /note= "primer rpo95"

FT primer\_bind 348..373

FT /tag= e

FT /note= "primer rpo105"

FT primer\_bind 354..373

FT /tag= f

FT misc\_feature /note= "primer KY290"  
 FT 372..373  
 FT /tag= g  
 FT /note= "M. tuberculosis signature nucleotide"  
 FT 433..434  
 FT /tag= h  
 FT /note= "M. tuberculosis signature nucleotide"  
 FT 438  
 FT /tag= i  
 FT /note= "M. tuberculosis signature nucleotide"  
 FT 468..469  
 FT /tag= j  
 FT /note= "M. tuberculosis signature nucleotide"  
 FT 486  
 FT /tag= k  
 FT /note= "M. tuberculosis signature nucleotide"  
 FT 501  
 FT /tag= l  
 FT /note= "M. tuberculosis signature nucleotide"  
 FT 516  
 FT /tag= m  
 FT /note= "M. tuberculosis signature nucleotide"  
 FT 516..535  
 FT /tag= n  
 FT /note= "primer rpo273"  
 FT 525  
 FT /tag= o  
 FT /note= "M. tuberculosis signature nucleotide"  
 FT 525..541  
 FT /tag= p  
 FT /note= "primer KY292"  
 FT 536..562  
 FT /tag= q  
 FT /note= "primer rpo293"  
 FT 640..666  
 FT /tag= r  
 FT /note= "primer rpo397"  
 FT 952..966  
 FT /tag= s  
 FT /note= "primer NMQRQ-1"  
 FT 952..966  
 FT /tag= t  
 FT /note= "primer NMQRQ-2"

W09533074-A1.

PD 07-DEC-1995.

XX 26-MAY-1995; 95WO-US06790.

XX 26-MAY-1994; 94US-0250030.

XX (HOFF ) HOFFMANN LA ROCHE INC.  
 PA (MAYO-) MAYO FOUNDATION.

XX Felmler TA, Hunt JM, Persing DH, Roberts GD, Whelen AC;  
 PI Young KKY;

XX WPI; 1996-030581/03.

FT Detection of Mycobacterium tuberculosis - by amplifying sample DNA  
 FT with a primer set that targets portions of the gene encoding rpoB.

XX Disclosure; Fig.3; 54pp; English.

XX This oligonucleotide DNA primer is specific for Mycobacterium  
 CC tuberculosis, and may be used to amplify a sample DNA by targeting  
 CC a portion of the gene encoding rpoB. The 1st several bases comprise a  
 CC nonhybridizing tail consisting of filler bases followed by  
 CC a restriction site incorporated to facilitate cloning using the  
 CC amplicon at a later date, if desired. The remaining bases hybridize  
 CC to bacterial rpoB DNA. The method provides for the detection of M.  
 CC tuberculosis and the concurrent determination of its drug



CC susceptibility, particularly to rifamycin. The method can provide  
CC often greater than 9% sensitivity and 100% specificity. The  
CC biological sample is a fluid or tissue sample from a human.  
XX  
SQ Sequence 970 BP; 182 A; 302 C; 330 G; 156 T; 0 other;  
Query Match 76.9%; Score 160; DB 17; Length 970;  
Best Local Similarity 85.6%; Pred. No. 3.2e-27;  
Matches 178; Conservative 0; Mismatches 30; Indels 0; Gaps 0;  
Qy 1 TCAAGGAGAAGCGCTACGACCTCCGCGGGTGGCGGTCTACAAGTCAACAGAAGCTCG 60  
Db 26 TCAAGGAGAAGCGCTACGACCTCCGCGGGTGGCGGTCTACAAGTCAACAGAAGCTCG 85  
Qy 61 GCCTGAACACCGCGTCCCGCATCAGCAGCACCACTCTGACCGAAGAGGACGCTCGGCCA 120  
Db 86 GCCTGCATGTGCGGGACCCATCAGTCTGCGACGTGACCGAAGAAGCTCGTGCCCA 145  
Qy 121 CCATCGAGTACCTGTCGCGCTCCAGAGGGCCACACACGATGACCGTCCCGGGCGGAG 180  
Db 146 CCATCGAATATCTGTCGCGCTTGCAGAGGTCAGACACGATGACCGTTCGCGGGCGCG 205  
Qy 181 TCGAGGTCCCGTGAACACGACAT 208  
Db 206 TCGAGGTCCCGTGAACACGACAT 233  
RESULT 14  
ID AAH51976 standard; DNA; 3519 BP.  
XX  
AC AAH51976;  
XX  
DT 04-SEP-2001 (first entry)  
XX  
DE Mycobacterium tuberculosis potential drug target gene SEQ ID 30.  
XX  
KW Drug target; growth; organism viability; characterisation; ds.  
XX  
OS Mycobacterium tuberculosis.  
XX  
PN WO200135317-A1.  
XX  
PD 17-MAY-2001.  
XX  
PF 13-NOV-2000; 2000WO-US31152.  
XX  
PR 12-NOV-1999; 99US-0165086.  
PR 12-NOV-1999; 99US-0165124.  
PR 01-FEB-2000; 2000US-0179531.  
XX  
PA (REGC ) UNIV CALIFORNIA.  
XX  
PI Eisenberg D, Rotstein SH, Marcotte EM;  
XX  
DR WPI: 2001-329193/34.  
DR P-PSDB; AAG81125.  
XX  
XX Identifying nucleotide or polypeptide sequence for use as drug target,  
PT involves providing algorithm that analyzes a functional relationship  
PT between nucleotide or polypeptide sequences, and comparing the  
PT sequences .  
XX  
PS Disclosure: Page 68-69; 207pp; English.  
XX  
XX This invention relates to a method for identifying a nucleotide or  
CC polypeptide sequence that may be a drug target, or essential for growth  
CC or viability of an organism. Polynucleotide sequences AAH51947 - AAH52092  
CC represent DNA encoding proteins AAG81096 - AAG81241, Mycobacterium  
CC tuberculosis proteins which are potential drug targets. The DNA and  
CC protein sequences are used to illustrate the method of the invention. The  
CC method involves providing an unknown nucleotide or polypeptide sequences,  
CC and comparing it to a number of sequences along with at least one

CC algorithm capable of analysing a functional relationship between  
CC nucleotide and polypeptide sequences. The method is useful for  
CC characterising the function of nucleic acids and polypeptides that may be  
CC useful as a target for a drug or essential for the growth or viability of  
CC an organism.  
XX  
SQ Sequence 3519 BP; 675 A; 1081 C; 1183 G; 580 T; 0 other;  
Query Match 76.9%; Score 160; DB 22; Length 3519;  
Best Local Similarity 85.6%; Pred. No. 3.2e-27;  
Matches 178; Conservative 0; Mismatches 30; Indels 0; Gaps 0;  
Qy 1 TCAAGGAGAAGCGCTACGACCTCCGCGGGTGGCGGTCTACAAGTCAACAGAAGCTCG 60  
Db 884 TCAAGGAGAAGCGCTACGACCTCCGCGGGTGGCGGTCTACAAGTCAACAGAAGCTCG 943  
Qy 61 GCCTGAACACCGCGTCCCGCATCAGCAGCACCACTCTGACCGAAGAGGACGCTCGGCCA 120  
Db 944 GCCTGCATGTGCGGGACCCATCAGTCTGCGACGTGACCGAAGAAGCTCGTGCCCA 1003  
Qy 121 CCATCGAGTACCTGTCGCGCTCCAGAGGGCCACACACGATGACCGTCCCGGGCGGAG 180  
Db 1004 CCATCGAATATCTGTCGCGCTTGCAGAGGTCAGACACGATGACCGTTCGCGGGCGCG 1063  
Qy 181 TCGAGGTCCCGTGAACACGACAT 208  
Db 1064 TCGAGGTCCCGTGAACACGACAT 1091  
RESULT 15  
ID AAH02079 standard; DNA; 3534 BP.  
XX  
AC AAH02079;  
XX  
DT 24-JUL-2001 (first entry)  
XX  
DE Mycobacterium tuberculosis nucleotide sequence SEQ ID NO:2072.  
XX  
KW Species specific; genus specific; family specific; probe; detection;  
KW identification; algal; archaeal; bacterial; fungal; parasitical;  
KW microorganism; diagnosis; translation elongation factor tu; toxin;  
KW translation elongation factor G; RecA recombinase; resistance;  
KW catalytic subunit of proton-translocating ATPase; antimicrobial;  
KW vaccine; primer; ds.  
XX  
OS Mycobacterium tuberculosis.  
XX  
PN WO200123604-A2.  
XX  
PD 05-APR-2001.  
XX  
PF 28-SEP-2000; 2000WO-CA01150.  
XX  
PR 28-SEP-1999; 99CA-2283458.  
PR 19-MAY-2000; 2000CA-2307010.  
XX  
PA (INFE-) INFECTIO DIAGNOSTIC (IDI) INC.  
XX  
PI Bergeron MG, Boissinot M, Huletsky A, Menard C, Ouellette M;  
PI Picard FJ, Roy PH;  
XX  
DR WPI: 2001-245006/25.  
XX  
PT Nucleic acid sequences are used to generate universal probes and  
PT primers which can be used to identify and detect the presence of algal,  
PT archaeal, bacterial, fungal and parasitical species in a test sample .  
XX  
PS Disclosure: Page 1478-1479; 1580pp; English.  
XX  
XX The present invention describes a method for generating a repository of  
CC nucleic acids of tuf, fus, atpD and/or recA genes from which probes  
CC and/or primers are derived. The method comprises amplifying the nucleic



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: November 12, 2002, 16:50:46 ; Search time 636.023 Seconds  
(without alignments)  
9517.553 Million cell updates/sec

Title: US-09-697-123B-17  
Perfect score: 208  
Sequence: 1 tcaaggagaagcgtacgac.....ccggtggaacacgacacat 208

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0  
Searched: 2054640 seqs, 14551402878 residues  
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*  
1: gb\_ba.\*  
2: gb\_htg.\*  
3: gb\_in.\*  
4: gb\_om.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_sts.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vi.\*  
15: em\_ba.\*  
16: em\_fun.\*  
17: em\_hum.\*  
18: em\_in.\*  
19: em\_mu.\*  
20: em\_om.\*  
21: em\_or.\*  
22: em\_ov.\*  
23: em\_pat.\*  
24: em\_ph.\*  
25: em\_pl.\*  
26: em\_ro.\*  
27: em\_sts.\*  
28: em\_un.\*  
29: em\_vi.\*  
30: em\_htg\_hum.\*  
31: em\_htg\_inv.\*  
32: em\_htg\_other.\*  
33: em\_htg\_mus.\*  
34: em\_htg\_pln.\*  
35: em\_htg\_rnd.\*  
36: em\_htg\_mam.\*  
37: em\_htg\_vrt.\*  
38: em\_sy.\*  
39: em\_htgo\_hum.\*  
40: em\_htgo\_mus.\*  
41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match %	Query Length	DB ID	Description
1	160	76.9	610	1	AJ318818 Mycobacte
2	160	76.9	610	1	MTU318819
3	160	76.9	616	1	AJ318813 Mycobacte
4	160	76.9	618	1	MTU318815
5	160	76.9	618	1	AJ318817 Mycobacte
6	160	76.9	633	1	MTU318814
7	160	76.9	637	1	MTU318816
8	160	76.9	639	1	AJ318821 Mycobacte
9	160	76.9	970	6	150706 Sequence 1
10	160	76.9	3534	6	AX111339 Sequence
11	160	76.9	3853	1	MTU12205
12	160	76.9	5084	1	MSGRPOB
13	160	76.9	19352	1	AE006964 Mycobacte
14	160	76.9	19770	1	MTCI376
15	152	73.1	3752	1	MSU24494
16	144	69.2	3447	6	AR067447 Sequence
17	144	69.2	37617	1	MLB1790G
18	144	69.2	348950	1	MLEPRIN7
19	117.2	56.3	3941	1	AF242549
20	83.6	40.2	32923	1	SCD82
21	81.6	39.2	3495	6	AX120631
22	81.6	39.2	328050	1	AP005275
23	81.6	39.2	34980	6	AX127144
24	49	23.6	18471	1	SCBAC17F8
25	47.8	23.0	1416	1	SRRECAGEN
26	47	22.6	195859	14	AF281817
27	46	22.1	33001	1	SCC123
28	45.6	21.9	7954	1	TAO19223
29	45.2	21.7	4575	1	AB083790
30	44.8	21.5	204050	1	AL646073
31	43.6	21.0	183046	8	AC087599
32	43.4	20.9	11103	1	AE004774
33	43.4	20.9	14210	1	AE012929
34	43.2	20.8	20054	1	SCCB12
35	42.8	20.6	2966	8	AF247649
36	42.8	20.6	23087	1	SCD40A
37	42.8	20.6	30753	1	SCC54
38	42.6	20.5	4333	1	AF012127
39	42.6	20.5	4556	1	TFSCSAB
40	42.4	20.4	12112	1	AE005722
41	42.4	20.4	30000	6	AX250262
42	42.2	20.3	3291	1	MSRECAGEN
43	42.2	20.3	7870	1	STE18574
44	42	20.2	204050	1	AL646070
45	41.8	20.1	1380	1	PSEARGFP

ALIGNMENTS

RESULT 1  
MTU318818  
LOCUS  
DEFINITION  
MTU318818 610 bp DNA circular BCT 09-AUG-2002  
Mycobacterium tuberculosis partial rpoB gene for RNA polymerase  
beta subunit, isolate I415-97.  
ACCESSION  
AJ318818  
VERSION  
AJ318818.1 GR:22208412  
KEYWORDS  
RNA polymerase beta subunit; rpoB gene.  
SOURCE  
Mycobacterium tuberculosis  
ORGANISM  
Mycobacterium tuberculosis  
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium  
tuberculosis complex.  
REFERENCE  
1  
Herrera, L., Jimenez, M.S. and Saez, J.A.

TITLE Molecular analysis of rifampin-resistant Mycobacterium tuberculosis isolated in Spain (1996-2001). Description of new alleles into rpoB gene and review

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 610)

AUTHORS Herrera,L.

TITLE Direct Submission

JOURNAL Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5, Majadahonda. Madrid. 28220, SPAIN

FEATURES

source 1..610

gene /organism="Mycobacterium tuberculosis"

CDS /isolate="1415-97"

1..610 /db\_xref="taxon:1773"

/gene="rpoB"

<1..>610

/gene="rpoB"

/codon\_start=1

/transl\_table=11

/product="RNA polymerase beta subunit"

/protein\_id="CAC87035.1"

/db\_xref="GI:22208413"

/translation="LDIYKLRPGEPPTKESAOITLLENLFKEKRYDLARVGRYKVK KLGLHVGEPITSLTEEDVATIEYLVRLHEGQTTMTVPGGVEVPVETDDIDHFGNR RLRTVGLIQNIQIRVGMSEMRVVRMTQDVEAITPOTLINIRPVVAAIKEFFGTS QPSQFMQNNPLSLGTHKRRLSALGPGLSRERAGLEVRDHP"

BASE COUNT 122 a 191 c 202 g 95 t

ORIGIN

Query Match 76.9%; Score 160; DB 1; Length 610;

Best Local Similarity 85.6%; Pred. No. 1e-17;

Matches 178; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

QY 1 TCAAGGAGAGCGCTACGACCTCGCGCGGTGGCGCTACAAGGTCAACAAGAGCTCG 60

Db 80 TCAAGGAGAGCGCTACGACCTCGCGCGGTGGCGCTACAAGGTCAACAAGAGCTCG 139

QY 61 GCCTGACACCGCTCGCGGATCAGCAGACCACTCTGACCGAAGAGAGCTCGTCGCCA 120

Db 140 GGCTGCATCTCGCGGAGCCCATCAGCTGTCGACGCTGACCGAAGAGAGCTCGTCGCCA 199

QY 121 CCATCGAGTACCTGTCGCGCTCGCGGATCAGCAGACCACTCTGACCGAAGAGAGCTCGTCGCCA 180

Db 200 CCATCGAATATCTGTCGCTTGCACGAGGTCACACCACTGACCGAAGAGAGCTCGTCGCCA 259

QY 181 TCGAGGTCCGCGTGGAAACCGACGACAT 208

Db 260 TCGAGGTCCGCGTGGAAACCGACGACAT 287

RESULT 2

MTU318819

LOCUS

DEFINITION

610 bp DNA circular BCT 09-AUG-2002

MTU318819 Mycobacterium tuberculosis partial rpoB gene for RNA polymerase beta subunit, isolate 1417-97.

ACCESSION

AJ318819

VERSION

AJ318819.1 GI:22208414

KEYWORDS

RNA polymerase beta subunit; rpoB gene.

SOURCE

Mycobacterium tuberculosis.

ORGANISM

Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium tuberculosis complex.

REFERENCE

1

AUTHORS

Herrera,L., Jimenez,M.S. and Saez,J.A.

TITLE

Molecular analysis of rifampin-resistant Mycobacterium tuberculosis isolated in Spain (1996-2001). Description of new alleles into rpoB gene and review

JOURNAL

Unpublished

REFERENCE

2 (bases 1 to 610)

AUTHORS

Herrera,L.

TITLE Direct Submission

JOURNAL Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5, Majadahonda. Madrid. 28220, SPAIN

FEATURES

source 1..610

gene /organism="Mycobacterium tuberculosis"

CDS /isolate="1417-97"

1..610 /db\_xref="taxon:1773"

/gene="rpoB"

<1..>610

/gene="rpoB"

/codon\_start=1

/transl\_table=11

/product="RNA polymerase beta subunit"

/protein\_id="CAC87036.1"

/db\_xref="GI:22208415"

/translation="LDIYKLRPGEPPTKESAOITLLENLFKEKRYDLARVGRYKVK KLGLHVGEPITSLTEEDVATIEYLVRLHEGQTTMTVPGGVEVPVETDDIDHFGNR RLRTVGLIQNIQIRVGMSEMRVVRMTQDVEAITPOTLINIRPVVAAIKEFFGTS QPSQFMQNNPLSLGTHKRRLSALGPGLSRERAGLEVRDHP"

BASE COUNT 122 a 191 c 202 g 95 t

ORIGIN

Query Match 76.9%; Score 160; DB 1; Length 610;

Best Local Similarity 85.6%; Pred. No. 1e-17;

Matches 178; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

QY 1 TCAAGGAGAGCGCTACGACCTCGCGCGGTGGCGCTACAAGGTCAACAAGAGCTCG 60

Db 80 TCAAGGAGAGCGCTACGACCTCGCGCGGTGGCGCTACAAGGTCAACAAGAGCTCG 139

QY 61 GCCTGAAACCGCGTCCCGCATCAGCAGACCACTCTGACCGAAGAGAGCTCGTCGCCA 120

Db 140 GGCTGCATCTCGCGGAGCCCATCAGCTGTCGACGCTGACCGAAGAGAGCTCGTCGCCA 199

QY 121 CCATCGAGTACCTGTCGCGCTCGCGGATCAGCAGACCACTGACCGAAGAGAGCTCGTCGCCA 180

Db 200 CCATCGAATATCTGTCGCTTGCACGAGGTCAGACCACTGACCGAAGAGAGCTCGTCGCCA 259

QY 181 TCGAGGTCCGCGTGGAAACCGACGACAT 208

Db 260 TCGAGGTCCGCGTGGAAACCGACGACAT 287

RESULT 3

MTU318813

LOCUS

DEFINITION

616 bp DNA circular BCT 09-AUG-2002

MTU318813 Mycobacterium tuberculosis partial rpoB gene for RNA polymerase beta subunit, isolate 1763-97.

ACCESSION

AJ318813

VERSION

AJ318813.1 GI:22208402

KEYWORDS

RNA polymerase beta subunit; rpoB gene.

SOURCE

Mycobacterium tuberculosis.

ORGANISM

Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium tuberculosis complex.

REFERENCE

1

AUTHORS

Herrera,L., Jimenez,M.S. and Saez,J.A.

TITLE

Molecular analysis of rifampin-resistant Mycobacterium tuberculosis isolated in Spain (1996-2001). Description of new alleles into rpoB gene and review

JOURNAL

Unpublished

REFERENCE

2 (bases 1 to 616)

AUTHORS

Herrera,L.

TITLE

Direct Submission

JOURNAL

Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5, Majadahonda. Madrid. 28220, SPAIN

FEATURES

source 1..616

```
/organism="Mycobacterium tuberculosis"
/isolate="1763-97"
/db_xref="taxon:1773"
1..615
/gene="rpoB"
<1..>615
/gene="rpoB"
/codon_start=1
/transl_table=11
/product="RNA polymerase beta subunit"
/protein_id="CAC87030.1"
/db_xref="GI:22208403"
translation="LDIYRKLRPGEPTTKESAQTLLLENLFFKRYDLARVGRYKVNK
KLGLHVEPITSTLTEDVDVATIEYLVRLHEGOTTMTVPGGVEVPVETDDIDHFGNR
RLRTVGEIQIRVGMRSRMRVVRERMTQDVEAITPTQLINIRPVAAIKEFFGTS
QLSQFIFMDONNPLSLGTHKRRLSALCPGLSRERAGLEVRDVP"
BASE COUNT 125 a 191 c 201 g 99 t
ORIGIN
Query Match 76.9%; Score 160; DB 1; Length 616;
Best Local Similarity 85.6%; Pred. No. 1e-17;
Matches 178; Conservative 0; Mismatches 30; Indels 0; Gaps 0;
Qy 1 TCAAGGAGAACGCTACGACCTCGCGGGTGGCGCGCTACAAAGGTCAACAAGAGCTCG 60
|||||
Db 80 TCAAGGAGAACGCTACGACCTCGCGGGTGGCGCGCTACAAAGGTCAACAAGAGCTCG 139
|||||
Qy 61 GCCTGAACACCGCTCCCGGATCAGCAGCACCACCTCTGACCGAGAGGAGCGTCGCGCCA 120
|||||
Db 140 GCCTGCGATGTGCGGAGCGCCATCAGCTCGTGCAGCGTGCAGGAGAGAGCGTCGCGCCA 199
|||||
Qy 121 CCATCGAGTACCTGCTCGCGCTGCACGAGGGCCACACGATGACCGTCCCGGGCGGAG 180
|||||
Db 200 CCATCGAATATCTGCTCGCTGCACGAGGTCAGACCCAGCATGACCGTTCGCGGGCGG 259
|||||
Qy 181 TCGAGGTGCGGTTGGAACCGACGACAT 208
|||||
Db 260 TCGAGGTGCGGTTGGAACCGACGACAT 287
|||||
RESULT 4
MTU318815 618 bp DNA circular BCT 09-AUG-2002
LOCUS
DEFINITION
Mycobacterium tuberculosis partial rpoB gene for RNA polymerase
beta subunit, isolate 2540-97.
ACCESSION
AJ318815
VERSION
AJ318815.1 GI:22208406
KEYWORDS
RNA polymerase beta subunit; rpoB gene.
SOURCE
Mycobacterium tuberculosis.
ORGANISM
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
tuberculosis complex.
REFERENCE
1
Herrera,L., Jimenez,M.S. and Saez,J.A.
Molecular analysis of rifampin-resistant Mycobacterium tuberculosis
isolated in Spain (1996-2001). Description of new alleles into rpoB
gene and review
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 618)
Herrera,L.
Direct Submission
Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro
Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5,
Majadahonda. Madrid. 28220, SPAIN
LOCATION/Qualifiers
1..618
/organism="Mycobacterium tuberculosis"
/isolate="2540-97"
/db_xref="taxon:1773"
1..618
/gene="rpoB"
<1..>618
gene
CDS
FEATURES
source
/organism="Mycobacterium tuberculosis"
/isolate="2540-97"
/db_xref="taxon:1773"
1..618
/gene="rpoB"
<1..>618
gene
CDS
```

```
/gene="rpoB"
/codon_start=1
/transl_table=11
/product="RNA polymerase beta subunit"
/protein_id="CAC87032.1"
/db_xref="GI:22208407"
translation="TDEALLDIYRKLRPGEPTTKESAOTLLLENLFFKRYDLARVGR
YKKNKGLHVGEPITSTLTEDVDVATIEYLVRLHEGOTTMTVPGGVEVPVETDDID
HFGNRLRTVGEIQIRVGMRSRMRVVRERMTQDVEAITPTQLINIRPVAAIKE
FFGTSQLSQFMDONNPLSLGTHKRRLSALCPGLSRERAGLEVRDVP"
BASE COUNT 124 a 190 c 207 g 97 t
ORIGIN
Query Match 76.9%; Score 160; DB 1; Length 618;
Best Local Similarity 85.6%; Pred. No. 1e-17;
Matches 178; Conservative 0; Mismatches 30; Indels 0; Gaps 0;
Qy 1 TCAAGGAGAACGCTACGACCTCGCGGGTGGCGCGCTACAAAGGTCAACAAGAGCTCG 60
|||||
Db 95 TCAAGGAGAACGCTACGACCTCGCGGGTGGCGCGCTACAAAGGTCAACAAGAGCTCG 154
|||||
Qy 61 GCCTGAACACCGCTCCCGGATCAGCAGCACCACCTCTGACCGAGAGGAGCGTCGCGCCA 120
|||||
Db 155 GCCTGCGATGTGCGGAGCGCCATCAGCTCGTGCAGCGTGCACGAGAGAGCGTCGCGCCA 214
|||||
Qy 121 CCATCGAGTACCTGCTCGCGCTGCACGAGGGCCACACGATGACCGTCCCGGGCGGAG 180
|||||
Db 215 CCATCGAATATCTGCTCGCTGCACGAGGTCAGACCCAGCATGACCGTTCGCGGGCGG 274
|||||
Qy 181 TCGAGGTGCGGTTGGAACCGACGACAT 208
|||||
Db 275 TCGAGGTGCGGTTGGAACCGACGACAT 302
|||||
RESULT 5
MTU318817 618 bp DNA circular BCT 09-AUG-2002
LOCUS
DEFINITION
Mycobacterium tuberculosis partial rpoB gene for RNA polymerase
beta subunit, isolate 2348-98.
ACCESSION
AJ318817
VERSION
AJ318817.1 GI:22208410
KEYWORDS
RNA polymerase beta subunit; rpoB gene.
SOURCE
Mycobacterium tuberculosis.
ORGANISM
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
tuberculosis complex.
REFERENCE
1
Herrera,L., Jimenez,M.S. and Saez,J.A.
Molecular analysis of rifampin-resistant Mycobacterium tuberculosis
isolated in Spain (1996-2001). Description of new alleles into rpoB
gene and review
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 618)
Herrera,L.
Direct Submission
Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro
Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5,
Majadahonda. Madrid. 28220, SPAIN
LOCATION/Qualifiers
1..618
/organism="Mycobacterium tuberculosis"
/isolate="2348-98"
/db_xref="taxon:1773"
1..618
/gene="rpoB"
<1..>618
/gene="rpoB"
/codon_start=1
/transl_table=11
/product="RNA polymerase beta subunit"
/protein_id="CAC87034.1"
/db_xref="GI:22208411"
gene
CDS
FEATURES
source
/organism="Mycobacterium tuberculosis"
/isolate="2348-98"
/db_xref="taxon:1773"
1..618
/gene="rpoB"
<1..>618
gene
CDS
```

/translation="TDEALLDIYRKLRPGEPTTKESAQTLLLENLFFKRYDLARVGR  
YKVNKKLGLHVGEPITSTLTEDVATIEYVLRHGGOTTMTVPGGVEVPVETDDID  
HFGNRLRTVGLIQNIIRVGMRSRMVRVRRMTTODVEAITPOTLINIRPVVAIKE  
FFGTSPQSFQMGQNNPLSLGTHKRLSALGPGGLSRERAGLEVRDV"  
BASE COUNT 124 a 192 c 207 g 95 t  
ORIGIN

Query Match 76.9%; Score 160; DB 1; Length 618;  
Best Local Similarity 85.6%; Pred. No. 1e-17;  
Matches 178; Conservative 0; Mismatches 30; Indels 0; Gaps 0;  
QY 1 TCAAGGAGAAGCGCTACGACCTCGCGCGGTGGCGCTACAAGGTCAACAAGAGCTCG 60  
DB 95 TCAAGGAGAAGCGCTACGACCTCGCGCGGTGGCGCTACAAGGTCAACAAGAGCTCG 154  
QY 61 GCCTGAACACCGCGTCCCGCATCAGCAGCACCTCTGACCGAAGAGGACGCTCGTCGCCA 120  
DB 155 GGCTGCATGTCGGCGAGGCCATCAGCTGTCGAGCGCTGACCGAAGAAGACGCTCGTCGCCA 214  
QY 121 CCATCGAGTACTGTGTCGCGCTGACGAGGGCCACACCATGACCGTCCCGGGCGGAG 180  
DB 215 CCATCGAATATCTGTCGCGCTTGACGAGGGTCAGACCATGACCGTTCCGGGCGGCG 274  
QY 181 TCGAGTCCCGGTGGAACCGACGACAT 208  
DB 275 TCGAGTCCCGGTGGAACCGACGACAT 302

RESULT 6  
MTU318814  
LOCUS  
DEFINITION Mycobacterium tuberculosis partial rpoB gene for RNA polymerase  
633 bp DNA circular BCT 09-AUG-2002  
beta subunit, isolate 1058-97.  
AJ318814  
ACCESSION  
VERSION AJ318814.1 GI:22208404  
KEYWORDS RNA polymerase beta subunit; rpoB gene.  
SOURCE Mycobacterium tuberculosis.  
ORGANISM Mycobacterium tuberculosis.  
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium  
tuberculosis complex.

REFERENCE 1  
AUTHORS Herrera, L., Jimenez, M.S. and Saez, J.A.  
TITLE Molecular analysis of rifampin-resistant Mycobacterium tuberculosis isolated in Spain (1996-2001). Description of new alleles into rpoB gene and review  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 633)  
AUTHORS Herrera, L.  
TITLE Direct Submission  
JOURNAL Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5, Majadahonda, Madrid. 28220, SPAIN

FEATURES  
source Location/Qualifiers  
1..633  
/organism="Mycobacterium tuberculosis"  
/isolate="1058-97"  
/db\_xref="taxon:1773"  
gene 1..633  
/gene="rpoB"  
CDS <1..>633  
/codon\_start=1  
/transl\_table=11  
/product="RNA polymerase beta subunit"  
/protein\_id="CAC87031.1"  
/db\_xref="GI:22208405"

translation="DNTVGTDEALLDIYRKLRPGEPTTKESAQTLLLENLFFKRYDL  
ARVGRYKVNKKLGLHVGEPITSTLTEDVATIEYVLRHGGOTTMTVPGGVEVPV  
TDDIDHFGNRLRTVGLIQNIIRVGMRSRMVRVRRMTTODVEAITPOTLINIRPVV  
AAIKFFGTSQLSQFMDQNNPLSLGTHKRLSALGPGGLSRERAGLEVRDV"  
BASE COUNT 129 a 195 c 210 g 99 t  
ORIGIN

Query Match 76.9%; Score 160; DB 1; Length 633;  
Best Local Similarity 85.6%; Pred. No. 1e-17;  
Matches 178; Conservative 0; Mismatches 30; Indels 0; Gaps 0;  
QY 1 TCAAGGAGAAGCGCTACGACCTCGCGCGGTGGCGCTACAAGGTCAACAAGAGCTCG 60  
DB 110 TCAAGGAGAAGCGCTACGACCTCGCGCGGTGGCGCTACAAGGTCAACAAGAGCTCG 169  
QY 61 GCCTGAACACCGCGTCCCGCATCAGCAGCACCTCTGACCGAAGAGGACGCTCGTCGCCA 120  
DB 170 GGCTGCATGTCGGCGAGGCCATCAGCTGTCGAGCGCTGACCGAAGAAGACGCTCGTCGCCA 229  
QY 121 CCATCGAGTACTGTGTCGCGCTGACGAGGGCCACACCATGACCGTCCCGGGCGGAG 180  
DB 230 CCATCGAATATCTGTCGCGCTTGACGAGGGTCAGACCATGACCGTTCCGGGCGGCG 289  
QY 181 TCGAGTCCCGGTGGAACCGACGACAT 208  
DB 290 TCGAGTCCCGGTGGAACCGACGACAT 317

RESULT 7  
MTU318816  
LOCUS  
DEFINITION Mycobacterium tuberculosis partial rpoB gene for RNA polymerase  
637 bp DNA circular BCT 09-AUG-2002  
beta subunit, isolate 1255-98.  
AJ318816  
ACCESSION  
VERSION AJ318816.1 GI:22208408  
KEYWORDS RNA polymerase beta subunit; rpoB gene.  
SOURCE Mycobacterium tuberculosis.  
ORGANISM Mycobacterium tuberculosis.  
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium  
tuberculosis complex.

REFERENCE 1  
AUTHORS Herrera, L., Jimenez, M.S. and Saez, J.A.  
TITLE Molecular analysis of rifampin-resistant Mycobacterium tuberculosis isolated in Spain (1996-2001). Description of new alleles into rpoB gene and review  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 637)  
AUTHORS Herrera, L.  
TITLE Direct Submission  
JOURNAL Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5, Majadahonda, Madrid. 28220, SPAIN

FEATURES  
source Location/Qualifiers  
1..637  
/organism="Mycobacterium tuberculosis"  
/isolate="1255-98"  
/db\_xref="taxon:1773"  
gene 1..637  
/gene="rpoB"  
CDS <1..>637  
/gene="rpoB"  
/codon\_start=1  
/transl\_table=11  
/product="RNA polymerase beta subunit"  
/protein\_id="CAC87033.1"  
/db\_xref="GI:22208409"

translation="VGTDEALLDIYRKLRPGEPTTKESAQTLLLENLFFKRYDLARV  
GRYKVNKKLGLHVGEPITSTLTEDVATIEYVLRHGGOTTMTVPGGVEVPVETDD  
IDHFGNRLRTVGLIQNIIRVGMRSRMVRVRRMTTODVEAITPOTLINIRPVVAI  
KEFFGTSQLSQFMDQNNPLSLGTHKRLSALGPGGLSRERAGLEVRDVHPSH"  
BASE COUNT 128 a 198 c 210 g 101 t  
ORIGIN  
Query Match 76.9%; Score 160; DB 1; Length 637;  
Best Local Similarity 85.6%; Pred. No. 1e-17;  
Matches 178; Conservative 0; Mismatches 30; Indels 0; Gaps 0;  
QY 1 TCAAGGAGAAGCGCTACGACCTCGCGCGGTGGCGCTACAAGGTCAACAAGAGCTCG 60

Db 101 TCAAGGAGAGCGCTACGACCTGCCCGCGTTCGCTGCTATTAAGGTCAACAAGAGCTCG 160  
Qy 61 GCCTGAACACCGCGTCCCGGATCAGACACACCTCTGACCGAAGAGAGCTCGTCGCCCA 120  
Db 161 GGCTGCATGTGTCGGCGGAGCCCATCAGTCGTCGACGCTGACCGAAGAAGAGCTCGTCGCCCA 220  
Qy 121 CCATCGAGTACCTGCTCGCGCTGCACGAGGCGCACACACGATGACCGTCCCGGCGGAG 180  
Db 221 CCATCGAATATCTGCTCGCGTTCGACGAGGCTGACACCGATGACCGTTCGCGGCGGCG 280  
Qy 181 TCGAGGTGCGCGGTGGAACCGAGACAT 208  
Db 281 TCGAGGTGCGCGGTGGAACCGAGACAT 308  
RESULT 8  
LOCUS MTU318821 639 bp DNA circular BCT 09-AUG-2002  
DEFINITION Mycobacterium tuberculosis partial rpoB gene for RNA polymerase  
ACCESSION AJ318821  
VERSION AJ318821.1 GI:22208418  
KEYWORDS RNA polymerase beta subunit; rpoB gene.  
SOURCE Mycobacterium tuberculosis.  
ORGANISM Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium  
tuberculosis complex.  
REFERENCE 1  
AUTHORS Herrera, L., Jimenez, M.S. and Saez, J.A.  
TITLE Molecular analysis of rifampin-resistant Mycobacterium tuberculosis  
isolated in Spain (1996-2001). Description of new alleles into rpoB  
gene and review  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 639)  
AUTHORS Herrera, L.  
TITLE Direct Submission  
JOURNAL Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro  
Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5,  
Majadahonda, Madrid: 28220, SPAIN  
FEATURES  
source  
1. .639  
/organism="Mycobacterium tuberculosis"  
/isolate="1071-98"  
/db\_xref="taxon:1773"  
gene 1. 639  
/gene="rpoB"  
CDS <1..>639  
/gene="rpoB"  
/codon\_start=1  
/trans\_table=11  
/product="RNA polymerase beta subunit"  
/protein\_id="CAC87038.1"  
/db\_xref="GI:22208419"  
/translation="RRTDEALLDIYKLRPGPEPTKESAOTLLENLFFKEKRYDLARV  
GRYKNKGLHVGEPITSSLTEDVVATIEYLVRLHGGOTTMTVPGVBPVETD  
IDHFNRLRTVGLIIONIRVMSRMRVVRMTQDVEAITPQTLINIRPAVAL  
KEFFCTOLSQFMVONPLSLGTLTKRRLSALPGSLSRERAGLEVRDHPHS"  
BASE COUNT 126 a 202 c 212 g 99 t  
ORIGIN  
Query Match 76.9%; Score 160; DB 1; Length 639;  
Best Local Similarity 85.6%; Pred. No. 1e-17;  
Matches 178; Conservative 0; Mismatches 30; Indels 0; Gaps 0;  
Qy 1 TCAAGGAGAGCGCTACGACCTGCCCGCGTTCGCTGCTATTAAGGTCAACAAGAGCTCG 160  
Db 101 TCAAGGAGAGCGCTACGACCTGCCCGCGTTCGCTGCTATTAAGGTCAACAAGAGCTCG 160  
Qy 61 GCCTGAACACCGCGTCCCGGATCAGACACACCTCTGACCGAAGAGAGCTCGTCGCCCA 120  
Db 161 GGCTGCATGTGTCGGCGGAGCCCATCAGTCGTCGACGCTGACCGAAGAAGAGCTCGTCGCCCA 220  
Qy 121 CCATCGAGTACCTGCTCGCGCTGCACGAGGCGCACACACGATGACCGTCCCGGCGGAG 180  
Db 221 CCATCGAATATCTGCTCGCGTTCGACGAGGCTGACACCGATGACCGTTCGCGGCGGCG 280  
Qy 181 TCGAGGTGCGCGGTGGAACCGAGACAT 208  
Db 281 TCGAGGTGCGCGGTGGAACCGAGACAT 308  
RESULT 9  
LOCUS 150706 970 bp DNA linear PAT 07-OCT-1997  
DEFINITION Sequence 1 from patent US 5643723.  
ACCESSION 150706  
VERSION 150706.1 GI:2472409  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 970)  
AUTHORS Persing, D.H., Hunt, J.J., Young, K.K.Y., Felmler, T.A., Roberts, G.D.  
and Whelan, A. Christian.  
TITLE Detection of a genetic locus encoding resistance to rifampin in  
mycobacterial cultures and in clinical specimens  
JOURNAL Patent: US 5643723-A 1 01-JUL-1997;  
FEATURES Location/Qualifiers  
source  
1. 970  
/organism="unknown"  
BASE COUNT 182 a 302 c 330 g 156 t  
ORIGIN  
Query Match 76.9%; Score 160; DB 6; Length 970;  
Best Local Similarity 85.6%; Pred. No. 9.7e-18;  
Matches 178; Conservative 0; Mismatches 30; Indels 0; Gaps 0;  
Qy 1 TCAAGGAGAGCGCTACGACCTGCCCGCGGTGGCGCGCTACAAGGTCAACAAGAGCTCG 60  
Db 26 TCAAGGAGAGCGCTACGACCTGCCCGCGGTGGCGCGCTACAAGGTCAACAAGAGCTCG 85  
Qy 61 GCCTGAACACCGCGTCCCGGATCAGACACGATGACCGTCTGACCGAAGAGAGCTCGTCGCCCA 120  
Db 86 GGCTGCATGTGCGGAGGCCCATCAGTGTGCGAGCTGACCGAAGAGAGCTCGTCGCCCA 145  
Qy 121 CCATCGAGTACCTGCTCGCGCTGCACGAGGCGCCACACGATGACCGTCCCGGCGGAG 180  
Db 146 CCATCGAATATCTGCTCGCGTTGCACGAGGCTGACACCGATGACCGTTCGCGGCGGCG 205  
Qy 181 TCGAGGTGCGCGGTGGAACCGAGACAT 208  
Db 206 TCGAGGTGCGCGGTGGAACCGAGACAT 233  
RESULT 10  
LOCUS AX111339 3534 bp DNA linear PAT 30-APR-2001  
DEFINITION Sequence 2072 from Patent WO0123604.  
ACCESSION AX111339  
VERSION AX111339.1 GI:13927631  
KEYWORDS  
SOURCE Mycobacterium tuberculosis.  
ORGANISM Mycobacterium tuberculosis.  
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
Actinomycetales; Corynebacterineae; Mycobacteriaceae;  
Mycobacterium; Mycobacterium tuberculosis complex.  
REFERENCE 1 (bases 1 to 3534)  
AUTHORS Bergeron, M.C., Boisset, M., Hueltgen, A., m Nard, C., Ouellette, M.,  
Picard, F.J. and Roy, P.H.  
TITLE Highly conserved genes and their use to generate probes and primers  
for detection of microorganisms  
JOURNAL Patent: WO 0123604-A 2072 05-APR-2001;  
FEATURES Infectio Diagnostic (I.D.I.) INC. (CA)  
Location/Qualifiers

Qy 121 CCATCGAGTACCTGCTCGCGCTGCACGAGGCGCACACGATGACCGTCCCGGCGGAG 180  
Db 221 CCATCGAATATCTGCTCGCGTTCGACGAGGCTGACACCGATGACCGTTCGCGGCGGCG 280  
Qy 181 TCGAGGTGCGCGGTGGAACCGAGACAT 208  
Db 281 TCGAGGTGCGCGGTGGAACCGAGACAT 308  
RESULT 9  
LOCUS 150706 970 bp DNA linear PAT 07-OCT-1997  
DEFINITION Sequence 1 from patent US 5643723.  
ACCESSION 150706  
VERSION 150706.1 GI:2472409  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 970)  
AUTHORS Persing, D.H., Hunt, J.J., Young, K.K.Y., Felmler, T.A., Roberts, G.D.  
and Whelan, A. Christian.  
TITLE Detection of a genetic locus encoding resistance to rifampin in  
mycobacterial cultures and in clinical specimens  
JOURNAL Patent: US 5643723-A 1 01-JUL-1997;  
FEATURES Location/Qualifiers  
source  
1. 970  
/organism="unknown"  
BASE COUNT 182 a 302 c 330 g 156 t  
ORIGIN  
Query Match 76.9%; Score 160; DB 6; Length 970;  
Best Local Similarity 85.6%; Pred. No. 9.7e-18;  
Matches 178; Conservative 0; Mismatches 30; Indels 0; Gaps 0;  
Qy 1 TCAAGGAGAGCGCTACGACCTGCCCGCGGTGGCGCGCTACAAGGTCAACAAGAGCTCG 60  
Db 26 TCAAGGAGAGCGCTACGACCTGCCCGCGGTGGCGCGCTACAAGGTCAACAAGAGCTCG 85  
Qy 61 GCCTGAACACCGCGTCCCGGATCAGACACGATGACCGTCTGACCGAAGAGAGCTCGTCGCCCA 120  
Db 86 GGCTGCATGTGCGGAGGCCCATCAGTGTGCGAGCTGACCGAAGAGAGCTCGTCGCCCA 145  
Qy 121 CCATCGAGTACCTGCTCGCGCTGCACGAGGCGCCACACGATGACCGTCCCGGCGGAG 180  
Db 146 CCATCGAATATCTGCTCGCGTTGCACGAGGCTGACACCGATGACCGTTCGCGGCGGCG 205  
Qy 181 TCGAGGTGCGCGGTGGAACCGAGACAT 208  
Db 206 TCGAGGTGCGCGGTGGAACCGAGACAT 233  
RESULT 10  
LOCUS AX111339 3534 bp DNA linear PAT 30-APR-2001  
DEFINITION Sequence 2072 from Patent WO0123604.  
ACCESSION AX111339  
VERSION AX111339.1 GI:13927631  
KEYWORDS  
SOURCE Mycobacterium tuberculosis.  
ORGANISM Mycobacterium tuberculosis.  
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
Actinomycetales; Corynebacterineae; Mycobacteriaceae;  
Mycobacterium; Mycobacterium tuberculosis complex.  
REFERENCE 1 (bases 1 to 3534)  
AUTHORS Bergeron, M.C., Boisset, M., Hueltgen, A., m Nard, C., Ouellette, M.,  
Picard, F.J. and Roy, P.H.  
TITLE Highly conserved genes and their use to generate probes and primers  
for detection of microorganisms  
JOURNAL Patent: WO 0123604-A 2072 05-APR-2001;  
FEATURES Infectio Diagnostic (I.D.I.) INC. (CA)  
Location/Qualifiers

source	1. .3534	/organism="Mycobacterium tuberculosis"	
		/strain="Rv"	
BASE COUNT	679 a	1081 c	1188 g 586 t
ORIGIN			
Query Match	76.9%;	Score 160;	DB 6; Length 3534;
Best Local Similarity	85.6%;	Pred. No. 7.7e-18;	
Matches 178;	Conservative 0;	Mismatches 30;	Indels 0; Gaps 0;
QY 1	TCAAGGAGAGCGGTACGACCTCGCGCGGTGGGCGGCTACAAAGTCAACAAAGAGCTCG	60	
Db	902 TCAAGGAGAGCGGTACGACCTCGCGCGGTGGGCGGCTACAAAGTCAACAAAGAGCTCG	961	
QY 61	GCCTGAACACCGGTCCCGGATCAGGACGACCTCTGACCGAAGAGAGCGTCTCGGCCA	120	
Db	962 GCCTGCATCTCGGAGCGCCATCAGTCTGACGCTGACCGAAGAGAGCGTCTCGGCCA	1021	
QY 121	CCATCGAGTACTGTGTCGCGCTGCACGAGGCGCACACCATGACCGTCCCGGGCGGAG	180	
Db	1022 CCATCGAATATCTGTCGCGCTTGCACGAGGTCAGACCGATGACCGTTCCTCGGGCGGC	1081	
QY 181	TCGAGGTCCCGTGAACACCGACGACAT	208	
Db	1082 TCGAGGTCCCGTGAACACCGACGACAT	1109	
RESULT 11			
LOCUS	MTU12205	3853 bp	DNA linear BCT 02-MAR-2000
DEFINITION	Mycobacterium tuberculosis H37Rv RNA-polymerase beta subunit (rpoB) gene, partial cds.		
ACCESSION	U12205		
VERSION	U12205.1	GI:515684	
KEYWORDS	Mycobacterium tuberculosis.		
SOURCE	Mycobacteria; Firmicutes; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium tuberculosis complex.		
REFERENCE	1 (bases 1 to 3853)		
AUTHORS	Imboden, P., Troller, R., Marchesi, F., Telenti, A., Bodmer, T., Cole, S., Schopfer, K. and Burkart, T.		
TITLE	The rpoB gene of Mycobacterium tuberculosis		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 3853)		
AUTHORS	Imboden, P.		
TITLE	Direct Submission		
JOURNAL	Submitted (11-JUL-1994) Paul Imboden, Institute for Medical Microbiology, University of Berne, Friedbuehlstrasse 51, Berne, 3010, Switzerland		
FEATURES	Location/Qualifiers		
source	1. .3853		
	/organism="Mycobacterium tuberculosis"		
	/strain="H37Rv"		
gene	576. .>3853		
	/db_xref="taxon:1773"		
CDS	576. .>3853		
	/gene="rpoB"		
	/gene="rpoB"		
	/codon_start=1		
	/transl_table=11		
	/product="RNA-polymerase beta subunit"		
	/protein_id="AAA20242.2"		
	/db_xref="GI:7144499"		
	/translation="M		
	REPLEVGLDVTDSFELWLGSPKRWRESAARGDVPVGLLEEVYELSPLEDFSGS		
	MSLSFSDPFDVDPKAPVDCKDMTYAAPLEVTAETFINNTGTGIKSTQVFMGDFPM		
	TEKGTPIINGTERRVVSQVSPGVFDETDKSTDKTLHSVKVIPSRGAWLEFDDVK		
	RTDVGRIIDRKRRPVTLLKALGWTSEQIVFSGFSEIMRSTLEKDNVTGTFDALLD		
	IYKLRPGEPPTKESAQTLLNLFKEKRYDLARVGRYKVNKKLGLHVGEPITSTLT		
	EEDVWATIEYLRVLRHEGQTTMTVPGVEVPVETDDIDHFGNRRRLRTVGLIQNQRV		

MSRMERVVRERTMTQDVEAITPOTLINIRPVVAAIKKEFFGTSQLSQFMDQNNPLSGLT			
HKRRLSALCPGGLSRRAGLEVDRVHPSHYGRMCPIETPEGNIGLIGLSLYIARVNP			
FGFIETPYRKVGVDSIGVYLTADIEDRHVVAOANSPIDAGRFVPRVLVRKAG			
EVEYVPSSEVDYMDVSPROMSVATAMIPFLEHDDANRALCANNQRAVPLVRSAP			
LVGTGMELRAIDAATSSQSSGVIEEVSADITIVHDNGTRRTYRMRKFAKSNHGT			
ANQCPIDAGDRVEAGVQADGPTDDGEMALGNLLVAIMPWGHNYDAIILSNRL			
DEEDVLTSHIEHEIDARDTKLGAEEITRDIPNISDEVLDLDERGTVIRIAGAEVRD			
VILGVKVTPTKETELTPEERLLRAIFGEKAREVRDTSKVPHGSEKVGIVRFSRED			
EDELPAQVNLVRYVVAOKRKISDGDKLGRHNGKVGIGKILPVEDMFLLEAQPNAIV			
ILNTHGVPRNRNIGIILETHLGCASHGKVDAAKGVDPMAARLPDELLEAQPNAIV			
TPVFGSAQEAELQGLLSCTLPNRDGDVLDADGKAMLEDFGSGSEPPFPYPTVGVYIM			
KLHLHLDKDIHARSTGPGVSMITQQPLGGKAQFGSGRFGEMECNAQATGAATTLQELL			
TKKS"			
BASE COUNT	723 a 1173 c 1293 g 664 t		
ORIGIN			
Query Match	76.9%;	Score 160;	DB 1; Length 3853;
Best Local Similarity	85.6%;	Pred. No. 7.5e-18;	
Matches 178;	Conservative 0;	Mismatches 30;	Indels 0; Gaps 0;
QY 1	TCAAGGAGAGCGGTACGACCTCGCGCGGTGGGCGGCTACAAAGTCAACAAAGAGCTCG	60	
Db	1477 TCAAGGAGAGCGGTACGACCTCGCGCGGTGGGCGGCTACAAAGTCAACAAAGAGCTCG	1536	
QY 61	GCCTGAACACCGGTCCCGGATCAGGACGACCTCTGACCGAAGAGAGCGTCTCGGCCA	120	
Db	1537 GCCTGCATCTCGGAGCGCCATCAGCTGTCGACGCTGACCGAAGAGAGCGTCTCGGCCA	1596	
QY 121	CCATCGAGTACTGTGTCGCGCTGCACGAGGCGCACACCATGACCGTCCCGGGCGGAG	180	
Db	1597 CCATCGAATATCTGTCGCGCTTGCACGAGGTCAGACCGATGACCGTTCCTCGGGCGGC	1656	
QY 181	TCGAGGTCCCGTGAACACCGACGACAT	208	
Db	1657 TCGAGGTCCCGTGAACACCGACGACAT	1684	
RESULT 12			
LOCUS	MSGRPOB	5084 bp	DNA linear BCT 13-SEP-1994
DEFINITION	Mycobacterium tuberculosis RNA polymerase beta-subunit (rpoB) gene, complete cds and RNA polymerase beta'-subunit rpoC gene, partial cds.		
ACCESSION	L27989.1	GI:468333	
VERSION	L27989.1	GI:468333	
KEYWORDS	RNA polymerase beta-subunit; rpoB gene.		
SOURCE	Mycobacterium tuberculosis (strain Rv) DNA.		
ORGANISM	Mycobacterium tuberculosis		
	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium tuberculosis complex.		
REFERENCE	1 (bases 1 to 5084)		
AUTHORS	Miller, L.P., Crawford, J.T. and Shinnick, T.M.		
TITLE	The rpoB gene of Mycobacterium tuberculosis		
JOURNAL	Antimicrob. Agents Chemother. 38 (4), 805-811 (1994)		
MEDLINE	94304130		
PUBMED	8031050		
FEATURES	Location/Qualifiers		
source	1. .5084		
	/organism="Mycobacterium tuberculosis"		
	/strain="Rv"		
gene	/db_xref="taxon:1773"		
	1065. .4598		
CDS	/gene="rpoB"		
	1065. .4598		
	/gene="rpoB"		
	/codon_start=1		
	/evidence=experimental		
	/transl_table=11		
	/product="RNA polymerase beta-subunit"		
	/protein_id="AAA21416.1"		
	/db_xref="GI:468334"		
	/translation="M		





```
SPASGRPLAMPRLDMVTGLYLLTTEVPDGTGEYQOPASGDHPETGVYSSPAEATMAADR
GVLVSRAKIKVRLTQTPRPVIEAEELFGHSGWOPGDAMWAETTLGRVMFNELLPLGYP
FVNKOMHKVQAAIINDLAERYPMIVVAQTVYDKLQAGFYWATRSQVYVSMADVLVPP
RKEILDHYERADKVEKQFQORGLNDRNEALVEITWKEATDEVQALREHYPPDDNP
IITIVDSGATQFQRTLAGMGLVYTNPKGEFIPRPVKSSEFREGVLTVEYFTNTHGA
RKLGLADTADSLGTLRRLVDVSDQVIVREHDCQTERGIVVELAERAPDGLTIRDP
YIETSAVARTLGDVADEAGNVIVRGDLDGDEIDALLAAGITQVKVRSVLTCAST
GVQATCYGRSMATKLVDIGEAVGIVAAQSIGEPGQTLMTRTPHOGSGVEDITGGIPLR
VQELFEARVPRGRAPLTADTVGRVLEDGERFKITIVPDDGGEVYVYDKLSRQRLRV
FKHEDSGSVLSODGHVEVQOQMEGSADPHEVLRVQGPREFVOIHLVREQVEYRAQG
VSHDXHIEVIVQMLKRVITIDSGSTFELPSGLIDRAEFEAENRRVVAEGGEPAAGR
PVLHGITKSLATSLMSAASFQETTRVLTDAAINCRSKLNGKLENKVIIGKLIIPAGT
GINRYRIAVQPTTEARAAAYTIPSYEDQYISPDFGAATCAAVPLDDYGSYDR"
/gene
CDS
complement(7691..8065)
/gene="MT0697"
complement(7691..8065)
/gene="MT0697"
/notes="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAK44923.1"
/db_xref="GI:13880220"
/translation="MFDSAAITNPGHAWASAMERSGLLECVCAGLDQBPFGFTADKL
NPORGSSRRVPRROADGGIATHVERGGQSQGGQAGVVPQRMHGFALAMQDRLIHH
GEQONRIQAQFRVFCVCSPT"
complement(8058..9972)
/gene="MT0698"
/notes="This region contains an authentic frame shift and
is not the result of a sequencing artifact; identified by
Glimmer2; putative; conserved hypothetical protein,
authentic frameshift"
10167..10925
/gene="MT0699"
10167..10925
/gene="MT0699"
/notes="identified by match to PFAM protein family HMM
PF01261"
/codon_start=1
/transl_table=11
/product="AP endonuclease, family 2"
/protein_id="AAK44924.1"
/db_xref="GI:13880221"
/translation="MLIGSHVSPTDPLAAAEAGADVQIFLGNPQSWKAPKPRDDAA
ALKAATLPIYHAPYLINLASANNRVRIPSRKILQETCAAAADIGAAAIVHGGHVD
DNDIDGFORWRKALDLETEVPVYLENTAGDHAMARRFDIARLWDVIGDTGIFGC
LDTCHTWAAGEALTDAVDRIKAITGRIDLHVCNDSRDEAGSGRDRHANLGSQIDPDL
LVAAVRAAGAPVTCETADQGRKDDIINFLERTGS"
10957..11799
/gene="MT0700"
10957..11799
/gene="MT0700"
/notes="similar to GB:U00012 PID:466863; identified by
sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="hydrolase/esterase, putative"
/protein_id="AAK44925.1"
/db_xref="GI:13880222"
/translation="MLRRVAILLAALAFAGCGGTRLAAGFGNGNSVHTLDVDCAGR
SVRLYKPVGLPSAPLVLMHGGEGSKAERSYGHDELASEKELVAYPDGYHRAWN
ANGGCGGPRAREGVDDIGFVRAVADIANNSIDFARVYVTGMSNGAIMSYTLACNT
SIFAAIGVSGTQDPCQSPRPVSVIHHGTADPLRVYHGGPGAFARIDGPPVPLN
AFMREYNGCVGLDITTEGPVTTSGATCADNRRRVLLTVDDAGHRWPSFATQTLWRFFA
AHFR"
11859..13487
/gene="MT0701"
11859..13487
/gene="MT0701"
/notes="similar to SP:P33224 GB:L20915 PID:457172
PID:457174 PID:537028; identified by sequence similarity;
putative"
/codon_start=1
```

```
/transl_table=11
/product="acyl-CoA dehydrogenase, putative"
/protein_id="AAK44926.1"
/db_xref="GI:13880223"
/translation="MSDTHVTVNOVPPLNKPASSPVLIEALIOEGOMGLDEVNEV
GAISCAOQWGLADNRRLIHTHDAYGYRDEVDYPAYHELMRTAITHGHHAAP
WADDPRAHYVRAAKTSWTEPCHICISMTYAVYPALRINSELAAYIEPLLTSREY
DEPLAPATITAGISMTQKQSGDVRAGTQATPNADGYSLTGHKMFPSAPMCD
IFVLQAQPDGSLFLLPRVLPDSTNRMFQRLKLDLGHANASSEVEYDGAVALV
GEGRGVPTIIEMVNLTLDCALASATSMRTGLTRAVHQAHRKAFGAYLIDOPLMRN
VLADLVAEAAATIVAMRAGATONAVRGNETEALLRRIGLAAAKYVCKRSTAAAE
ALECLGGNGYVEDSCMPRLYREAPLWGTWEGSGNVSALDTLRANATRPACVEVLFDEL
ARSAGNCR2LDGSHVERLPLQGLDITIGYRARKTAEDICLALQGLSLVRHCHPAVEA
FLATVLLCGWGGANGTTPAGLDLAPILERALVKG"
/gene
CDS
13498..14436
/gene="MT0702"
13498..14436
/gene="MT0702"
/notes="similar to GP:3885480; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="enoyl-CoA hydratase/isomerase family protein"
/protein_id="AAK44927.1"
/db_xref="GI:13880224"
/translation="MTHAIRPYDFDNLKTMWYTVTGRIARITFNRPEKGNALIAIDTPL
ELSLVERADLDGPHVILVSGRGEFCAGFDLSAYAGSSSTGGGAYOGTVLDGKT
QAVNHLPNQMDPMIDYOMSRFVRGFSALMHADKPTVKTHGYCVAGGTDLALHADQ
VTAADAKIGWPPTRVMGVPAAGLWHRFLGDQRAKRLIFTGDCITGAQAAMGLAVEA
PEPADLDERLVARIAALPVNOLIMVKLALNSALLQOQVATSRMVSTVFDGAARHT
PEGHAFVADAVEHGFDRVRRDEPFQDYGQASRV"
14439..15161
/gene="MT0703"
14439..15161
/gene="MT0703"
/notes="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
Query Match 76.9%; Score 160; DB 1; Length 19352;
Best Local Similarity 85.6%; Pred. No. 5.6e-18;
Matches 178; Conservative 0; Mismatches 30; Indels 0; Gaps 0;
QY 1 TCAAGGAGAAGCGCTACGACCTCGCGCGGTGCGGCTACAGGTCAACAAGAGCTCG 60
DB 1064 TCAAGGAGAAGCGCTACGACCTCGCGCGGTGCGGCTACAGGTCAACAAGAGCTCG 1123
QY 61 GCCTGAACACCGCGTCCCGCATCAGCAGCACCTCTGACCGAAGAGCGCTCGCGCA 120
DB 1124 GCCTGATGTCGCGGAGCCATCAGCTCGTCCGACCTGACCGAAGAGCGCTCGCGCA 1183
QY 121 CCATCGAGTACCTGGTCCGCTGACGAGGCGCCACACCATGATACCGCGGCGGAG 180
DB 1184 CCATCGAATATCTGTCGCTTGCACGAGGCTCAGACCATGATACCGCGGCGGCGG 1243
QY 181 TCGAGGTGCGGTGGAACCGAGCAT 208
DB 1244 TCGAGGTGCGGTGGAACCGAGCAT 1271
RESULT 14
MTCI376
LOCUS
DEFINITION
Mycobacterium tuberculosis H37Rv.
ACCESSION
295972 ALI23456
VERSION
295972.1 GI:3261790
KEYWORDS
Mycobacterium tuberculosis H37Rv.
ORGANISM
Mycobacterium tuberculosis H37Rv.
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Corynebacterineae; Mycobacteriaceae;
Mycobacterium; Mycobacterium tuberculosis complex.
```

REFERENCE  
AUTHORS

1 (bases 1 to 19770)  
Cole, S.T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C.,  
Harris, D., Gordon, S.V., Eiglmeier, K., Gas, S., Barry III, C.E.,  
Tekala, F., Badcock, K., Basham, D., Brown, D., Chillingworth, T.,  
Connor, R., Davies, R., Devlin, K., Feltwell, T., Gentles, S.,  
Hamlin, N., Holroyd, S., Hornsby, T., Jagels, K., Krogh, A., McLean, J.,  
Moule, S., Murphy, L., Oliver, S., Osborne, J., Quail, M.A.,  
Rajandream, M.A., Rogers, J., Rutter, S., Seeger, K., Skellon, S.,  
Squares, S., Squires, R., Sulston, J.E., Taylor, K., Whitehead, S. and  
Barrell, B.G.  
Deciphering the biology of Mycobacterium tuberculosis from the  
complete genome sequence  
Nature 393 (6685), 537-544 (1998)  
98295987  
PUBMED  
9634230  
2 (bases 1 to 19770)  
Parkhill, J.  
Direct Submission  
Submitted (11-JUN-1998) Submitted on behalf of the Mycobacterium  
tuberculosis sequencing and mapping teams: Sanger Centre, Wellcome  
Trust Genome Campus, Hinxton, Cambridge CB10 1SA Unite de Genetique  
Moleculaire Bacterienne, Institut Pasteur, 28 rue du Docteur Roux,  
75724 Paris Cedex 15, France E-mail: parkhill@sanger.ac.uk  
On Jun 27, 1998 this sequence version replaced gi:2143285.

## COMMENT

Notes:  
Details of M. tuberculosis sequencing at the Sanger Centre are  
available on the World Wide Web.  
(URL, [http://www.sanger.ac.uk/Projects/M\\_tuberculosis/](http://www.sanger.ac.uk/Projects/M_tuberculosis/)) CDS have  
been renumbered from the original cosmid submissions but the old  
gene designations are in brackets after the new gene numbers.  
Gene prediction was based on a Hidden Markov Model of TB genes  
implemented in TBparse (Krogh) supplemented with visual inspection  
of positional base preference in codons, especially where there is  
an increase in the observed/expected third position G + C.  
CAUTION: In some cases we may not have predicted the correct  
initiation codon. Where possible we choose an initiation codon  
(atg, gtg, or ttg) which is preceded by an upstream ribosome  
binding site sequence (optimally 5-13bp before the initiation  
codon). If this cannot be identified we choose the most upstream  
initiation codon.

FEATURES  
source

Location/Qualifiers  
1. .19770  
/organism="Mycobacterium tuberculosis H37Rv"  
/strain="H37Rv"  
/db\_xref="taxon:83332"  
/clone="1376"  
57. .61  
/note="agga, possible rbs upstream of Rv0654"  
68. .1573  
/gene="Rv0654"  
68. .1573  
/gene="Rv0654"  
/note="Rv0654"  
/note="Rv0654, (MTC1376.22), len: 501. unknown, FASTA  
score: Q53353 LIGNOSTILBENE-ALPHA, BETA-DIOXYGENASE (485  
aa) opt:280 z-score: 330.1 E(): 2.3e-11, (28.5% identity  
in 523 aa overlap). Also similar to M. tuberculosis protein  
MTCY21C12.07c (29.5% identity in 522 aa overlap)"  
/codon\_start=1  
/transl\_table=11  
/product="hypothetical protein Rv0654"  
/protein\_id="CAB09380.1"  
/db\_xref="GI:2143307"  
/db\_xref="SPTREMBL:O06785"  
/translation="MTTAAAESQNPYLEGLAPVSTVATDLPVTGRIPHELDGRY  
LRGNPNPAEVPDPATVHFTGDAMVGVALRGDKARWRNRRVTPAACAALGPISA  
RHPRTGIIEGPNLVNTHAGTLVLAEGVGNHETDELDTGVCDFDGLTGGYT  
YVYIDLPTFPMQVPASVPRLQRPALVQSVLGRVRLPDPAAALGNRMQGHSD  
RLPYAWNPSPYRQVYMPREGNEQVRFDIIEPCYVYHPFNAYSECRNGAEVLVDVV  
RYSRMFDRRGPGGDRPSLDRTWTLNLTAGTAVTAECRDRAQEPRLNLTLCGPHR  
FATVIGIEGFLVGAGAAALSTPLYKQDCVTSSTVATVLPQRPVPMGFHGNWAPT"  
DGILMGYWHHGRGEGQLLLDQAQTLSTIATVHLPPQRPVPMGFHGNWAPT"  
1585. .2664

## gene

## CDS

/gene="Rv0655"  
1585. .2664  
/gene="Rv0655"  
/note="Rv0655, (MTC1376.21), len: 359, abc transporter,  
FASTA score: YRBF\_ECOU1 P45393 hypothetical abc  
transporter atp-binding (269 aa) opt: 644 z-score: 721.8  
E(): 3.4e-33 (38.5% identity in 244 aa overlap): contains  
PS00017 ATP/GTP-binding site motif A, PS00211 ABC  
transporters family signature, highly similar to M. leprae  
MKL\_MYCLE P30769 possible ribonucleotide transport atp-  
(347 aa) opt: 2021 z-score:2244.4 E(): 0, (92.2% identity  
in 335 aa overlap). Also similar to many other M.  
tuberculosis ABC transporters eg. MTCY253.24 (33.6%  
identity in 241 aa overlap)"  
/codon\_start=1  
/transl\_table=11  
/product="hypothetical protein Rv0655"  
/protein\_id="CAB09379.1"  
/db\_xref="GI:2143306"  
/db\_xref="SWISS-PROT:O06784"  
/translation="MRYSDSYHTTGRWQPRASTEGFPMGVSIEVNGLTKSFSSRIWE  
DVTLTIPAGEVSVLLGPGSGTKGVFLKSLIGLLRPERGSIIDGTDIIKCSAKELYEI  
RTLFGVLQDQKALFGSMNLNDYNTAPPLREHTPKKKESEIRDIYMEKALVGLGGDEKRF  
PGEISGGMRKRAGLARVLDPDIIILCDPEPSGLDPVATYLSQLMDINAOIDATIL  
IVYHINARTVPDNGMLFRRLHVMFGPREVLLTSDPEVYVQFLNGRRIRGPIGNSEE  
KDEATMAEEQALLDAGHHAGGVETEGYVPPQISATPGMPERKAVARQARVREMLHTL  
PKKAAAILDDLEGTHTKYAVHEIGQ"  
1762. .1785  
/gene="Rv0655"  
/note="PS00017 ATP/GTP-binding site motif A"  
2074. .2118  
/gene="Rv0655"  
/note="PS00211 ABC transporters family signature"  
complement(3052. .3435)  
/gene="Rv0656c"  
complement(3052. .3435)  
/gene="Rv0656c"  
/note="Rv0656c, (MTC1376.20, unknown), len: 127 aa"  
/codon\_start=1  
/transl\_table=11  
/product="hypothetical protein Rv0656c"  
/protein\_id="CAB09378.1"  
/db\_xref="GI:2143305"  
/db\_xref="SPTREMBL:O06783"  
/translation="MAAATTTGTHRGLELRAAQRAVGSCEPQRAEFCRSARNADEFDQ  
MSRMFGDVPDPVPKSVWRWIDSAQHRLARAGAVGALSVVDLLICDTAAAGLVLHL  
DDADYELAERHLPDIRVRVVSADD"  
complement(3530. .3685)  
/gene="Rv0657c"  
complement(3530. .3685)  
/gene="Rv0657c"  
/note="Rv0657c, (MTC1376.19), unknown, len: 51 aa; similar  
to several other M. tuberculosis hypothetical proteins eg  
YW08 MYCTU Q10848 hypothetical 8.9 kd protein cy39.08c (80  
aa) fasta scores, opt: 107 z-score: 182.3 E(): 0.0038,  
45.8% identity in 48 aa overlap. Also similar to MTCY48\_5  
and AL020958|SC4H8\_7 Streptomyces coelicolor cosmid 4H8 (66  
aa), 41.0% identity in 39 aa overlap."  
/codon\_start=1  
/transl\_table=11  
/product="hypothetical protein Rv0657c"  
/protein\_id="CAB09377.1"  
/db\_xref="GI:2143304"  
/db\_xref="SPTREMBL:O06782"  
/translation="MSVTQIDLDDEALADVMRIAIVHTKKEAVNLAMRDYVERFRRIE  
ALARSRE"  
complement(3691. .3696)  
/note="possible RBS upstream of Rv0657c"  
complement(3761. .4477)  
/gene="Rv0658c"  
complement(3761. .4477)  
/gene="Rv0658c"  
/note="Rv0658c, (MTC1376.18), len: 238, unknown,

## RBS

complement(3691. .3696)  
/note="possible RBS upstream of Rv0657c"  
complement(3761. .4477)  
/gene="Rv0658c"  
complement(3761. .4477)  
/gene="Rv0658c"  
/note="Rv0658c, (MTC1376.18), len: 238, unknown,

## gene

complement(3691. .3696)  
/note="possible RBS upstream of Rv0657c"  
complement(3761. .4477)  
/gene="Rv0658c"  
complement(3761. .4477)  
/gene="Rv0658c"  
/note="Rv0658c, (MTC1376.18), len: 238, unknown,

## CDS

complement(3691. .3696)  
/note="possible RBS upstream of Rv0657c"  
complement(3761. .4477)  
/gene="Rv0658c"  
complement(3761. .4477)  
/gene="Rv0658c"  
/note="Rv0658c, (MTC1376.18), len: 238, unknown,



Db 1248 TCGAGTTCGCCGTCGAGTCGCACGACAT 1275

Search completed: November 13, 2002, 01:26:47  
Job time : 660.723 secs

GenCore version 5.1.3  
Copyright (c) 1993 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 12, 2002, 16:59:42 : Search time 1189.12 Seconds  
(without alignments)  
3037.202 Million cell updates/sec

Title: US-09-697-123b-13

Perfect score: 223

Sequence: 1 tcaaggagaagcgctacgac.....ccggtggaaccgacgacat 223

Scoring table: IDENTITY\_NUC

Gapop 10.0, Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: em\_estba:\*

2: em\_esthum:\*

3: em\_estin:\*

4: em\_estnu:\*

5: em\_estov:\*

6: em\_estpl:\*

7: em\_estro:\*

8: em\_htc:\*

9: gb\_est1:\*

10: gb\_est2:\*

11: gb\_htc:\*

12: gb\_est3:\*

13: gb\_est4:\*

14: gb\_est5:\*

15: em\_estfun:\*

16: em\_estom:\*

17: gb\_gss:\*

18: em\_gss\_hum:\*

19: em\_gss\_inv:\*

20: em\_gss\_pin:\*

21: em\_gss\_vrt:\*

22: em\_gss\_fun:\*

23: em\_gss\_mam:\*

24: em\_gss\_mus:\*

25: em\_gss\_other:\*

26: em\_gss\_pro:\*

27: em\_gss\_rod:\*

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	123.8	55.5	1282	9	AI770311
2	42.4	19.0	431	13	BI097733
3	42.4	19.0	507	13	BI096671
C 4	42.4	19.0	531	13	BM348454
C 5	42.4	19.0	579	13	BM340507
6	42.4	19.0	586	12	BG842655

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

C 7	42.4	19.0	593	13	BM079854
C 8	42.4	19.0	597	13	BM333485
C 9	42.4	19.0	606	13	BM337568
C 10	42.4	19.0	610	13	BM073467
C 11	42.4	19.0	613	13	BM340956
C 12	42.4	19.0	622	13	BM333757
C 13	42.4	19.0	651	13	BM333944
C 14	42.4	19.0	659	13	BM266851
C 15	42.4	19.0	659	13	BM337566
C 16	42.4	19.0	687	13	BM349198
C 17	42.4	19.0	701	13	BM075264
C 18	41.6	18.7	530	13	BM336791
C 19	40.4	18.1	511	10	AV628007
20	40.4	18.1	543	10	AV631761
C 21	40.4	18.1	554	10	AV631497
C 22	40.4	18.1	725	14	B0823567
C 23	40.2	18.0	524	13	BM336777
C 24	40	17.9	546	12	BF257512
C 25	39.4	17.7	494	13	BJ318566
C 26	39.4	17.7	649	13	BJ312832
C 27	39.2	17.6	433	10	AM923980
C 28	39.2	17.6	531	12	BG841942
C 29	39.2	17.6	536	10	AM679799
C 30	39.2	17.6	546	10	AV435406
C 31	39.2	17.6	604	13	BI721593
C 32	39.2	17.6	655	9	AU170926
C 33	39	17.5	300	9	AU176504
C 34	39	17.5	363	9	AU162829
C 35	38.8	17.4	632	13	BM266831
C 36	38.4	17.2	562	14	B0752716
C 37	38	17.0	459	9	AU033206
C 38	38	17.0	461	14	D49085
C 39	37.8	17.0	533	10	BE405789
C 40	37.8	17.0	566	13	BM076409
C 41	37.8	17.0	594	13	BJ317721
C 42	37.8	17.0	617	13	BJ323305
C 43	37.8	17.0	710	12	BG321153
C 44	37.8	17.0	1057	12	BF345788
C 45	37.8	17.0	1856	11	AY109418

## ALIGNMENTS

RESULT 1	AI770311/c
LOCUS	AI770311
DEFINITION	42 Mycobacterium anaerobic stationary phase library Mycobacterium smegmatis CDNA, mRNA sequence.
ACCESSION	AI770311
VERSION	AI770311.1
KEYWORDS	GI:6742680
SOURCE	EST.
ORGANISM	Mycobacterium smegmatis.
REFERENCE	Bacteria: Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
AUTHORS	1 (bases 1 to 1282)
TITLE	Murugasu-Oei, B., Tay, A. and Dick, T.
JOURNAL	upregulation of stress response genes and ABC transporters in anaerobic stationary-phase Mycobacterium smegmatis
MEDLINE	Mol. Gen. Genet. 262 (4-5), 677-682 (1999)
COMMENT	20092472

Contact: Murugasu-Oei, B.  
Mycobacterium Laboratory  
Institute of Molecular and Cell Biology  
30 Medical Drive, Singapore 117609, Republic of Singapore  
Tel: 65 874 3011  
Fax: 65 779 1117  
Email: mcbom@imcb.nus.edu.sg  
Insert Length: 1282 Std Error: 0.00  
Seq primer: T3 Forward; T7 Backward.  
Location/Qualifiers  
1. 1282







REFERENCE	Clade; Panicoideae; Andropogoneae; Zea.
AUTHORS	Wen, T.J., Qiu, F., Qiu, F., Guo, L., Ashlock, D.A and Schnable, P.S.
TITLE	Expressed Sequence Tags from B73 Maize: various stages and tissues including seedlings treated with a variety of hormones
JOURNAL	unpublished (2001)
COMMENT	Contact: Patrick S. Schnable Schnable Laboratory Iowa State University G405 Agronomy, Iowa State University, Ames, IA 50011-1010, USA Tel: 515-294-0975 Fax: 515-294-2299 Email: schnable@iastate.edu Individual basecall and confidence value were assigned using the Phred software, ( <a href="http://depts.washington.edu/ventures/collabtr/direct/index.htm#fb">http://depts.washington.edu/ventures/collabtr/direct/index.htm#fb</a> rt). Overall sequence quality assessment and vector trimming were conducted using the Lucy software ( <a href="http://www.tigr.org/soflab/">http://www.tigr.org/soflab/</a> ). Lucy parameters were set to ensure an overall trimmed quality of 97.5% or better without any vector fragments in the chosen high-quality region of each sequence. Low-quality bases between the poly-T and the high-quality region were replaced with N's to serve as spacers. PCR Primers FORWARD: primer T7-1 (AA TAC GAC TCA CTA TAG) BACKWARD: primer T3 (ATT AAC CCT CAC TAA AG) Seq primer: primer T3 (ATT AAC CCT CAC TAA AG). Location/Qualifiers 1..579
FEATURES	
source	

```

1. 373
/organism="Zea mays"
/cultivar="B73"
/db_xref="taxon:4577"
/clone="MEST323-F04"
/clone_lib="ISUM5-RN"
/tissue_type="mixed"
/lab_host="DH108"
/notice=vector: pT73PAC; Site_1: EcoRI; Site_2: NotI;
Tissues: Germinated seed and seedlings (1, 2, 8, 11 DAG),
Mixed mature tissues (17, 21, 38, 69, 77 DAG), kernels
(3, 5, 10, 15, 20, 25, 30, 36 DAG), Adventitious roots (65 DAG
(3, 5, 10, 15, 20, 25, 30, 36 DAG), Immature ear (0.2-3.0
cm, 53, 56, 59 DAG), Husk (73 DAG), Silk, unpollinated
first ear, ear shank, etiolated seedlings, callus
Cycloneximide-treated callus, Anaerobic treated seedlings
NAA (a-Naphthalene acetic acid)-treated seedlings,
Kinetin-treated seedlings, ACPc
(1-aminocyclopropane-1-carboxylic acid)-treated seedlings,
Brassinolide-treated seedlings, ABA (Abscissic acid
)-treated seedlings, GA (Gibberellic acid)-treated
seedlings, JA (Jasmonic acid)-treated seedlings, ds-cDNA
molecules were generated as follows. First-strand cDNA was
prepared from oligo-dT selected mRNA by priming with a
NotI oligo-dT primer (5'
AATGTGAAGAAATTCGGCGCCAGGAAATTTTTTTTTTTT). The
resulting DNA:RNA hybrid was treated with RNase H and used
as a template for DNA PolI-catalyzed second strand
synthesis. After the addition of EcoRI adaptors, the
ds-cDNAs were digested with NotI and size-selected. The
resulting molecules were directionally cloned into the
EcoRI and NotI sites of the pT73PAC vector. The library
then went through one round of normalization to cDNA value
of 5 based on the methods of Marcello Bento Soares (Genome
Research 6: 791-806, 1996)."
123 a 145 c 147 g 163 t 1 others

```

	Query Match	19.08;	Score 42.4;	DB 13;	Length 579;
	Best Local Similarity	58.9%;	Pred. No. 2.5;		
	Matches 73;	Conservative	0;	Mismatches 51;	Indels 0;
	Gaps	0;			
Qy	4	AGGAGAACGCTAGCACTGTCGCCCGCTCGCCGCTACAAAGTCAACAAAGCTGGGCC	63		
Db	509	AGCTGATCGCGCAGCACTGCGCACCGATCGCGGCAAGGAGTACCAAGAAAGGTCGGCC	450		

Qy	64	TGATACACCGATATCCGATCACCAACACGCGCTGACCGAAGAAGACGTGTCGCCACCA	123
Db	449	TCGTCTACCAAGCTCAACATCGCCCCCAAGAAGATCGCGTTGATGAGAGATCTTCGTGG	390
Qy	124	TCGA	127
Db	389	TCGA	386
RESULT	6		
LOCUS	BG842655		
DEFINITION	MEST34-E06.T7-1 ISUM3-TL Zea mays cDNA clone MEST34-E06 5', mRNA		
ACCESSION	BG842655	586 bp	linear
VERSION	BG842655.2	GI:14244717	
KEYWORDS	EST.		
SOURCE	Zea mays.		
ORGANISM	Zea mays		
REFERENCE	Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.		
AUTHORS	Qiu,F., Cui,F., Guo,L., Ashlock,D.A., Wen,T.J. and Schnable,P.S.		
TITLE	Expressed Sequence Tags from B73 Maize Seedlings and Silks		
JOURNAL	Unpublished (2001)		
COMMENT	On May 25, 2001 this sequence version replaced gi:14208977.		
	Contact: Patrick S. Schnable		
	Schnable Laboratory		
	Iowa State University		
	G405 Agronomy, Iowa State University, Ames, IA 50011-1010, USA		
	Tel: 515-294-0975		
	Fax: 515-294-2299		
	Email: schnable@iastate.edu		
	PCR Primers		
	FORWARD: T7-1 (AA TAC GAC TCA CTA TAG)		
	BACKWARD: T3 (ATT AAC CCT CAC TAA AG)		
	Seq primer: primer T7-1 (AA TAC GAC TCA CTA TAG).		

```

FEATURES             source
seq primer: primer: 17-1 (AA TAC GAC TCA CTA TAG).
Location/Qualifiers
1. .586
/organism="Zea mays"
/cultivar="B73"
/db_xref="taxon:4577"
/clone="MEST34-E06"
/clone_lib="ISUM3-TL"
/tissue_type="Seedling and silk"
/lab_host="DH10B"
/note="Vector: pT73PAC; Site_1: EcoRI; Site_2: NotI;
ds-cDNA molecules were generated as follows. First-strand
cDNA was prepared from oligo-dT selected mRNA by priming
with a NotI oligo-dT primer (5'
AAGTCGAACAAATTCGGCGCCGAGGAATTTTTTTTTTTTTT). The
resulting DNA:RNA hybrid was treated with RNase H and used
as a template for DNA PolI-catalyzed second strand
synthesis. After the addition of EcoRI adaptors, the
ds-cDNAs were digested with NotI and size-selected. The
resulting molecules were directionally cloned into the
EcoRI and NotI sites of the pT73PAC vector."
148 a 1172 c 143 q 123 t
BASE COUNT

```

[illegible]



```

PCR Primers
  FORWARD: primer T7-1 (AA TAC GAC TCA CTA TAG)
  BACKWARD: primer T3 (ATT AAC CCT CAC TAA AG)
Seq primer: primer T3 (ATT AAC CCT CAC TAA AG).

FEATURES
  source
    1. .606
      /organism="Zea mays"
      /cultivar="B73"
      /db_xref="taxon:4577"
      /clone="WEST209-B07"
      /clone_lib="ISUM5-RN"
      /tissue_type="mixed"
      /lab_host="DH10B"
      /note="Vector: pT73PAC; Site_1: EcoRI; Site_2: NotI;
Tissues: Germinated seed and seedlings (1, 2, 8, 11 DAG),
Mixed mature tissues (17, 21, 38, 69, 77 DAG), Kernels
(3, 5, 10, 15, 20, 25, 30, DAP), Adventitious roots (65 DAG
), Tassel (3-39 cm, 53 and 56 DAG), Immature ear (0.2-3.0
cm, 53, 56, 59 DAG), Husk (73 DAG), Silk, unpollinated
first ear, ear shank, etiolated seedlings, callus,
Cycloheximide-treated callus, Anaerobic treated seedlings
, NAA (a-Naphthalene acetic acid)-treated seedlings,
Kinetin-treated seedlings, ACPC

```

Cycloheximide-treated callus, Anaerobic treated seedlings, NAA (a-Naphthalene acetic acid)-treated seedlings, Kinetin-treated seedlings, ACPIC (1-aminocyclopropane-1-carboxylic acid)-treated seedlings, Brassinolide-treated seedlings, ABA (Abscissic acid)-treated seedlings, GA (Gibberellic acid)-treated seedlings, JA (Jasmonic acid)-treated seedlings. ds-cDNA molecules were generated as follows. First-strand cDNA was prepared from oligo-dT selected mRNA by priming with a NotI oligo-dT primer (5' AACTGGAAGAAATTCGGCGCGCAATTTTTTTTTTTTTT). The resulting DNA:RNA hybrid was treated with RNase H and used as a template for DNA PolI-catalyzed second strand synthesis. After the addition of EcoRI adaptors, the ds-cDNAs were digested with NotI and size-selected. The resulting molecules were directionally cloned into the EcoRI and NotI sites of the pT73PAC vector. The library then went through one round of normalization to Cot value of 5 based on the methods of Marcelo Bento Soares (Genome Research 6: 791-806, 1996)."

129 a	150 c	156 g	158 t	13 others
-------	-------	-------	-------	-----------

BASE COUNT  
ORIGIN

BASE COUNT    129 a    150 c    156 g    158 t    13 others  
 ORIGIN  
 Query Match                    19.0%;    Score 42.4;    DB 13;    Length 606;  
 Best Local Similarity    58.9%;    Pred. No. 2.5;  
 Matches 73; Conservative    0; Mismatches 51; Indels 0; Gaps 0;

QY	64	TGAACACCGGATCATCCGGATCACCACACGAGCGTGCACCGAAGAGACGTCGTCGCCACCA	123
Db	448	TCGTCTACCAAGCTCAACATCGCCCCCAAGAAGATCGCGTTGATGAGGAGATCTTCGTGG	389
QY	124	TCGA	127
Db	388	TCGA	385

RESULT	10
LOCUS	BM073467/c
DEFINITION	MEST67-H11.T3 ISU4-TN Zea mays cDNA clone MEST67-H11 3' , mRNA
ACCESSION	BM073467
VERSION	BM073467.1
KEYWORDS	GI:16917872
SOURCE	EST.
ORGANISM	Zea mays. Zea mays. Zea mays.
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.









GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: November 12, 2002, 16:50:46 : Search time 636.023 Seconds  
(without alignments)  
9517.553 Million cell updates/sec

Title: US-09-697-123b-8  
Perfect score: 208  
Sequence: 1 tcaagagaagcgctacgac.....ccggtgaaaccgacgacat 208

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues  
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :		GenEmbl:
1:	gb_ba:	*
2:	gb_hgt:	*
3:	gb_in:	*
4:	gb_om:	*
5:	gb_ov:	*
6:	gb_pat:	*
7:	gb_ph:	*
8:	gb_pl:	*
9:	gb_pr:	*
10:	gb_ro:	*
11:	gb_sts:	*
12:	gb_sy:	*
13:	gb_un:	*
14:	gb_vl:	*
15:	em_ba:	*
16:	em_fun:	*
17:	em_hum:	*
18:	em_in:	*
19:	em_mu:	*
20:	em_om:	*
21:	em_or:	*
22:	em_ov:	*
23:	em_pat:	*
24:	em_ph:	*
25:	em_pl:	*
26:	em_ro:	*
27:	em_sts:	*
28:	em_un:	*
29:	em_vi:	*
30:	em_htg_hum:	*
31:	em_htg_inv:	*
32:	em_htg_other:	*
33:	em_htg_mus:	*
34:	em_htg_pln:	*
35:	em_htg_rod:	*
36:	em_htg_mam:	*
37:	em_htg_vrt:	*
38:	em_sy:	*
39:	em_htgo_hum:	*
40:	em_htgo_mus:	*
41:	em_htgo_other:	*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	166.4	80.0	610	1	MTU318818	AJ318818 Mycobacte
2	166.4	80.0	610	1	MTU318819	AJ318819 Mycobacte
3	166.4	80.0	616	1	MTU318813	AJ318813 Mycobacte
4	166.4	80.0	618	1	MTU318815	AJ318815 Mycobacte
5	166.4	80.0	618	1	MTU318817	AJ318817 Mycobacte
6	166.4	80.0	633	1	MTU318814	AJ318814 Mycobacte
7	166.4	80.0	637	1	MTU318816	AJ318816 Mycobacte
8	166.4	80.0	639	1	MTU318821	AJ318821 Mycobacte
9	166.4	80.0	970	6	150706	150706 Sequence 1
10	166.4	80.0	3534	6	AX111339	AX111339 Sequence
11	166.4	80.0	3853	1	MTU12205	U12205 Mycobacteri
12	166.4	80.0	5084	1	MSGRPOB	L27989 Mycobacteri
13	166.4	80.0	19352	1	AE006964	AE006964 Mycobacte
14	166.4	80.0	19770	1	MTC1376	295972 Mycobacteri
15	156.8	75.4	3752	1	MSU24494	U24494 Mycobacteri
16	142.4	68.5	3447	6	AR067447	AR067447 Sequence
17	142.4	68.5	37617	1	MLB1790G	Z14314 M.leprae ge
18	142.4	68.5	348950	1	MLEPRTN7	AL583923 Mycobacte
19	118.8	57.1	3941	1	AF242549	AF242549 Mycolato
20	80.4	38.7	32923	1	SCD82	AL160431 Streptomy
21	80	38.5	3495	6	AX120631	AX120631 Sequence
22	80	38.5	328050	1	AP005275	AP005275 Corynebac
23	80	38.5	349980	6	AX127144	AX127144 Sequence
24	47.2	22.7	139467	1	D90905	D90905 Synchocyst
25	46.6	22.4	204050	1	AL646070	AL646070 Ralstonia
26	45.8	22.0	204050	1	AL646073	AL646073 Ralstonia
27	45.8	22.0	347660	1	AP002994	AP002994 Mesorhizo
28	43.2	20.8	3027	3	AY061587	AY061587 Drosophil
29	43.2	20.8	3920	3	AY119444	AY119444 Drosophil
30	43.2	20.8	4470	3	AF001796	AF001796 Drosophil
31	43.2	20.8	125766	2	ACU13137	ACU13137 Drosophil
32	43.2	20.8	171376	3	AC007929	AC007929 Drosophil
33	43.2	20.8	220842	3	AE003745	AE003745 Drosophil
34	43	20.7	186159	3	AC008202	AC008202 Drosophil
35	42.6	20.5	18471	1	SCBAC17F8	AL596030 Streptomy
36	41.6	20.0	349116	1	AP003003	AP003003 Mesorhizo
37	41.4	19.9	1246	1	AF469953	AF469953 Streptomy
38	41.2	19.8	152203	2	OSJN00166	AF62967 Oryza sat
39	41	19.7	195859	14	AF281817	AF281817 Tupaia he
40	40.8	19.6	204050	1	AL646073	AL646073 Ralstonia
41	40.6	19.5	7870	1	STE18574	Y18574 Streptomyce
42	40.6	19.5	14210	1	AE012929	AE012929 Chlorobiu
43	40.4	19.4	4556	1	TFSCSAB	X54073 T.flavus sc
44	40.2	19.3	1242	6	E13276	E13276 Tamias asia
45	40.2	19.3	1367	10	AB000545	AB000545 Tamias s1

ALIGNMENTS

RESULT 1	MTU318818	MTU318818	610 bp	DNA	circular BCT 09-AUG-2002
LOCUS	Mycobacterium tuberculosis partial rpoB gene for RNA polymerase				
DEFINITION	beta subunit, isolate 1415-97.				
ACCESSION	AJ318818				
VERSION	AJ318818.1	GI:22208412			
KEYWORDS	RNA polymerase beta subunit; rpoB gene.				
SOURCE	Mycobacterium tuberculosis.				
ORGANISM	Mycobacterium tuberculosis				
REFERENCE	Bacteria: Actinobacteria; Actinobacteridae; Actinomycetales;				
AUTHORS	Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium tuberculosis complex.				
	Herrera, L., Jimenez, M.S. and Saez, J.A.				



TITLE Molecular analysis of rifampin-resistant Mycobacterium tuberculosis isolated in Spain (1996-2001). Description of new alleles into rpoB gene and review

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 610)

AUTHORS Herrera, L.

TITLE Direct Submission

JOURNAL Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5, Majadahonda. Madrid. 28220, SPAIN

FEATURES

source 1. .610

gene /organism="Mycobacterium tuberculosis"

CDS /isolate="1415-97"

gene /db\_xref="taxon:1773"

gene /gene="rpoB"

CDS <1..>610

gene /gene="rpoB"

gene /codon\_start=1

gene /transl\_table=11

gene /product="RNA polymerase beta subunit"

gene /protein\_id="CAC87035.1"

gene /db\_xref="GI:22208413"

BASE COUNT 122 a 191 c 202 g 95 t

ORIGIN

Query Match 80.0%; Score 166.4; DB 1; Length 610;

Best Local Similarity 87.5%; Pred. No. 1.1e-21;

Matches 182; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 1 TCAAGGAGAACGCTAGCAGCTGGCCGCTGTCGGCCGATACAAAGTCAACGAAGCTGG 60

Db 80 TCAAGGAGAACGCTAGCAGCTGGCCGCTGTCGGCCGATACAAAGTCAACGAAGCTGG 139

QY 61 GCCTGAACACCAATCATCCGATCACCACGACGCTGACCGAAGAGACGCTCGCGCCA 120

Db 140 GGCTGCATGTCGGGAGCCCATCACGTCGACGCTGACCGAAGAGACGCTCGCGCCA 199

QY 121 CCATCAGATATCTGTCGGCTGCACGAGGCCGACGATGACCGTGCCTGGCGGGG 180

Db 200 CCATCGAATATCTGTCGGCTGCACGAGGTCAGACGACGATGACCGTTCCTGGCGGGG 259

QY 181 TCCAGGTGCGGTGGAAACCGACGACAT 208

Db 260 TCGAGTGGCGGTGGAAACCGACGACAT 287

RESULT 2

MTU318819

LOCUS Mycobacterium tuberculosis partial rpoB gene for RNA polymerase

DEFINITION beta subunit, isolate 1417-97.

ACCESSION AJ318819

VERSION 1 GI:22208414

KEYWORDS RNA polymerase beta subunit; rpoB gene.

SOURCE Mycobacterium tuberculosis.

ORGANISM Mycobacterium tuberculosis.

REFERENCE 1

AUTHORS Herrera, L., Jimenez, M.S. and Saez, J.A.

TITLE Molecular analysis of rifampin-resistant Mycobacterium tuberculosis isolated in Spain (1996-2001). Description of new alleles into rpoB gene and review

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 610)

AUTHORS Herrera, L.

TITLE Direct Submission

JOURNAL Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5, Majadahonda. Madrid. 28220, SPAIN

FEATURES

source 1. .610

gene /organism="Mycobacterium tuberculosis"

CDS /isolate="1417-97"

gene /db\_xref="taxon:1773"

gene /gene="rpoB"

CDS <1..>610

gene /gene="rpoB"

gene /codon\_start=1

gene /transl\_table=11

gene /product="RNA polymerase beta subunit"

gene /protein\_id="CAC87036.1"

gene /db\_xref="GI:22208415"

BASE COUNT 122 a 191 c 202 g 95 t

ORIGIN

Query Match 80.0%; Score 166.4; DB 1; Length 610;

Best Local Similarity 87.5%; Pred. No. 1.1e-21;

Matches 182; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 1 TCAAGGAGAACGCTAGCAGCTGGCCGCTGTCGGCCGATACAAAGTCAACGAAGCTGG 60

Db 80 TCAAGGAGAACGCTAGCAGCTGGCCGCTGTCGGCCGATACAAAGTCAACGAAGCTGG 139

QY 61 GCCTGAACACCAATCATCCGATCACCACGACGCTGACCGAAGAGACGCTCGCGCCA 120

Db 140 GGCTGCATGTCGGGAGCCCATCACGTCGACGCTGACCGAAGAGACGCTCGCGCCA 199

QY 121 CCATCAGATATCTGTCGGCTGCACGAGGCCGACGATGACCGTGCCTGGCGGGG 180

Db 200 CCATCGAATATCTGTCGGCTGCACGAGGTCAGACGACGATGACCGTTCCTGGCGGGG 259

QY 181 TCGAGGTGCGGTGGAAACCGACGACAT 208

Db 260 TCGAGTGGCGGTGGAAACCGACGACAT 287

RESULT 3

MTU318813

LOCUS Mycobacterium tuberculosis partial rpoB gene for RNA polymerase

DEFINITION beta subunit, isolate 1763-97.

ACCESSION AJ318813

VERSION 1 GI:22208402

KEYWORDS RNA polymerase beta subunit; rpoB gene.

SOURCE Mycobacterium tuberculosis.

ORGANISM Mycobacterium tuberculosis.

REFERENCE 1

AUTHORS Herrera, L., Jimenez, M.S. and Saez, J.A.

TITLE Molecular analysis of rifampin-resistant Mycobacterium tuberculosis isolated in Spain (1996-2001). Description of new alleles into rpoB gene and review

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 616)

AUTHORS Herrera, L.

TITLE Direct Submission

JOURNAL Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5, Majadahonda. Madrid. 28220, SPAIN

FEATURES

source 1. .616

/organism="Mycobacterium tuberculosis"  
/isolate="1763-97"  
/db\_xref="taxon:1773"  
1..615  
/gene="rpoB"  
/gene="rpoB"  
/codon\_start=1  
/transl\_table=11  
/product="RNA polymerase beta subunit"  
/protein\_id="CAC87030.1"  
/db\_xref="GI:22208403"  
/translation="LDIYRKLRPGEPTTKESQTLLENLFFKRYDLARVGRYKVNK  
KLGHLGEPITSTLFEEDVATIEYLRHGGQTTMTVPGGVEVPVETDDIDHFGNR  
RLTVGELIQNIRVGMRSRMRVRRMTQDVEAITPOTLINIRPVAAIKEFFGTS  
OLSOQIFMDQNNPLSLGTHKRRLSALGPGLSRERAGLEVRDVH"

BASE COUNT 125 a 191 c 201 g 99 t  
ORIGIN  
Query Match 80.0%; Score 166.4; DB 1; Length 616;  
Best Local Similarity 87.5%; Pred. No. 1.1e-21;  
Matches 182; Conservative 0; Mismatches 26; Indels 0; Gaps 0;  
QY 1 TCAAGGAGAGCGCTACGACCTGGCGGTGTCGGCGGATACAGAGTCAACAAGAGCTGG 60  
Db 80 TCAAGGAGAGCGCTACGACCTGGCGGTGTCGGCGGATACAGAGTCAACAAGAGCTGG 139  
QY 61 GCCTGAACACCAATCATCCGATCACCAGCAGCGCTGACCGAAGAGAGCGTGGCGCCA 120  
Db 140 GCCTGCGATGTCGGGAGCCCATCAGCTGCTGCGAGCTGACCGAAGAGAGCGTGGCGCCA 199  
QY 121 CCATCGAGTATCTGTCGGCTGCACGAGGGCCAGCCGATGACCGTCCGGGGGGGG 180  
Db 200 CCATCGAATATCTGTCGGCTGTCAGGAGGTCAGACGATGACCGTCCGGGGGGGG 259  
QY 181 TCGAGGTGCGGTGGAACCGAGCAGCAT 208  
Db 260 TCGAGGTGCGGTGGAACCGAGCAGCAT 287

RESULT 4  
MTU318815  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
source  
gene  
CDS

MTU318815 618 bp DNA circular BCT 09-AUG-2002  
Mycobacterium tuberculosis partial rpoB gene for RNA polymerase  
beta subunit, isolate 2540-97.  
AJ318815  
RNA polymerase beta subunit; rpoB gene.  
Mycobacterium tuberculosis.  
Mycobacterium tuberculosis.  
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium  
tuberculosis complex.  
1  
Herrera, L., Jimenez, M.S. and Saez, J.A.  
Molecular analysis of rifampin-resistant Mycobacterium tuberculosis  
isolated in Spain (1996-2001). Description of new alleles into rpoB  
gene and review  
Unpublished  
2 (bases 1 to 618)  
Herrera, L.  
Direct Submission  
Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro  
Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5,  
Majadahonda. Madrid. 28220, SPAIN  
Location/Qualifiers  
1..618  
/organism="Mycobacterium tuberculosis"  
/isolate="2540-97"  
/db\_xref="taxon:1773"  
1..618  
/gene="rpoB"  
<1..>618

/gene="rpoB"  
/codon\_start=1  
/transl\_table=11  
/product="RNA polymerase beta subunit"  
/protein\_id="CAC87032.1"  
/db\_xref="GI:22208407"  
/translation="TDEALLDIYRKLRPGEPTTKESQTLLENLFFKRYDLARVGR  
YKVNKKLGHVGEPTSTLFEEDVATIEYLRHGGQTTMTVPGGVEVPVETDDID  
HFGNRLRTVGLIQNIRVGMRSRMRVRRMTQDVEAITPOTLINIRPVAAIKE  
FFGTSQLSQFMDQNNPLSLGTHKRRLSALGPGLSRERAGLEVRDV"

BASE COUNT 124 a 190 c 207 g 97 t  
ORIGIN  
Query Match 80.0%; Score 166.4; DB 1; Length 618;  
Best Local Similarity 87.5%; Pred. No. 1.1e-21;  
Matches 182; Conservative 0; Mismatches 26; Indels 0; Gaps 0;  
QY 1 TCAAGGAGAGCGCTACGACCTGGCGGTGTCGGCGGATACAGAGTCAACAAGAGCTGG 60  
Db 95 TCAAGGAGAGCGCTACGACCTGGCGGTGTCGGCGGATACAGAGTCAACAAGAGCTGG 154  
QY 61 GCCTGAACACCAATCATCCGATCACCAGCAGCGCTGACCGAAGAGAGCGTGGCGCCA 120  
Db 155 GCCTGCGATGTCGGGAGCCCATCAGCTGCTGCGAGCTGACCGAAGAGAGCGTGGCGCCA 214  
QY 121 CCATCGAGTATCTGTCGGCTGCACGAGGGCCAGCCGATGACCGTCCGGGGGGGG 180  
Db 215 CCATCGAATATCTGTCGGCTTGCAGGAGGTCAGACGATGACCGTCCGGGGGGGG 274  
QY 181 TCGAGGTGCGGTGGAACCGAGCAGCAT 208  
Db 275 TCGAGGTGCGGTGGAACCGAGCAGCAT 302

RESULT 5  
MTU318817  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
source  
gene  
CDS

MTU318817 618 bp DNA circular BCT 09-AUG-2002  
Mycobacterium tuberculosis partial rpoB gene for RNA polymerase  
beta subunit, isolate 2348-98.  
AJ318817  
RNA polymerase beta subunit; rpoB gene.  
Mycobacterium tuberculosis.  
Mycobacterium tuberculosis.  
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium  
tuberculosis complex.  
1  
Herrera, L., Jimenez, M.S. and Saez, J.A.  
Molecular analysis of rifampin-resistant Mycobacterium tuberculosis  
isolated in Spain (1996-2001). Description of new alleles into rpoB  
gene and review  
Unpublished  
2 (bases 1 to 618)  
Herrera, L.  
Direct Submission  
Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro  
Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5,  
Majadahonda. Madrid. 28220, SPAIN  
Location/Qualifiers  
1..618  
/organism="Mycobacterium tuberculosis"  
/isolate="2348-98"  
/db\_xref="taxon:1773"  
1..618  
/gene="rpoB"  
<1..>618  
/codon\_start=1  
/transl\_table=11  
/product="RNA polymerase beta subunit"  
/protein\_id="CAC87034.1"  
/db\_xref="GI:22208411"

/translation="TTEALLDIYRKLRPGEPTKESAOQLLENLFFKEKRYDLARVGR  
YKNNKGLHVGEPITSTLTEDVATIEYLRHHEGOTMTVPVGVPEVPTDDID  
HFGNRRRTVGLIQNIRVGMRSRMRVVRMTTQDVEAITPOTLINIRPVVAIKE  
FFGTQSQFMGQNNPLSLTKRRLSALPGGLSRERAGLEVRDV"  
BASE COUNT 124 a 192 c 207 g 95 t

Query Match 80.0%; Score 166.4; DB 1; Length 618;  
Best Local Similarity 87.5%; Pred. No. 1.1e-21;  
Matches 182; Conservative 0; Mismatches 26; Indels 0; Gaps 0;  
ORIGIN  
QY 1 TCAAGGAGAGCGCTACGACCTGGCCGCTGTCGCCGATACAAAGGTCACAAAGAGCTGG 60  
Db 95 TCAAGGAGAGCGCTACGACCTGGCCGCTGTCGCCGATACAAAGGTCACAAAGAGCTGG 154  
QY 61 GCCTGAACACCAATCCGATCCACGAGAGCGCTGACCGAAGAAGACGCTCGTCGCCA 120  
Db 155 GGCTGCATGTCGGCGAGCCCATCAGCTCGTCAGCGTGACCGAAGAAGACGCTCGTCGCCA 214  
QY 121 CCATCGAGTATCTGTCGCCCTGACGAGGCGCCAGGCGCCAGCATGACCGTGGCGGGGG 180  
Db 215 CCATCGAATATCTGTCGCCCTGTCAGGAGGTCAGACACGATGACCGTTCCGGGGCGCG 274  
QY 181 TCGAGGTCCGGTGGAAACCGACGACAT 208  
Db 275 TCGAGGTCCGGTGGAAACCGACGACAT 302

RESULT 6  
MTU318814  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
source  
gene  
CDS

MTU318814 633 bp DNA circular BCT 09-AUG-2002  
Mycobacterium tuberculosis partial rpoB gene for RNA polymerase  
beta subunit, isolate 1058-97.  
AJ318814  
AJ318814.1 GI:22208404  
RNA polymerase beta subunit; rpoB gene.  
Mycobacterium tuberculosis.  
Mycobacterium tuberculosis.  
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium  
tuberculosis complex.  
Herrera, L., Jimenez, M.S. and Saez, J.A.  
Molecular analysis of rifampin-resistant Mycobacterium tuberculosis  
isolated in Spain (1996-2001). Description of new alleles into rpoB  
gene and review  
Unpublished  
2 (bases 1 to 633)  
Herrera, L.  
Direct Submission  
Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro  
Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5,  
Majadahonda, Madrid. 28220, SPAIN

Location/Qualifiers  
1..633  
/organism="Mycobacterium tuberculosis"  
/isolate="1058-97"  
/db\_xref="taxon:1773"  
1..633  
/gene="rpoB"  
<1..>633  
/codon\_start=1  
/transl\_table=11  
/product="RNA polymerase beta subunit"  
/protein\_id="CAC87031.1"  
/db\_xref="GI:22208405"

/translation="DNTVGTDEALLDIYRKLPGSPPTKESAOQLLENLFFKEKRYDL  
ARVGRYKNNKGLHVGEPITSTLTEDVATIEYLRHHEGOTMTVPVGVPEVPTDD  
TDDIDHFGNRRRTVGLIQNIRVGMRSRMRVVRMTTQDVEAITPOTLINIRPVV  
AAIKFFGTQSQFMGQNNPLSLTKRRLSALPGGLSRERAGLEVRDV"  
BASE COUNT 129 a 195 c 210 g 99 t

Query Match 80.0%; Score 166.4; DB 1; Length 633;  
Best Local Similarity 87.5%; Pred. No. 1.1e-21;  
Matches 182; Conservative 0; Mismatches 26; Indels 0; Gaps 0;  
ORIGIN  
QY 1 TCAAGGAGAGCGCTACGACCTGGCCGCTGTCGCCGATACAAAGGTCACAAAGAGCTGG 60  
Db 110 TCAAGGAGAGCGCTACGACCTGGCCGCTGTCGCCGATACAAAGGTCACAAAGAGCTGG 169  
QY 61 GCCTGAACACCAATCCGATCCACGAGAGCGCTGACCGAAGAAGACGCTCGTCGCCA 120  
Db 170 GGCTGCATGTCGGCGAGCCCATCAGCTCGTCAGCGTGACCGAAGAAGACGCTCGTCGCCA 229  
QY 121 CCATCGAGTATCTGTCGCCCTGACGAGGCGCCAGGCGCCAGCATGACCGTGGCGGGGG 180  
Db 230 CCATCGAATATCTGTCGCCCTGTCAGGAGGTCAGACACGATGACCGTTCCGGGGCGCG 289  
QY 181 TCGAGGTCCGGTGGAAACCGACGACAT 208  
Db 290 TCGAGGTCCGGTGGAAACCGACGACAT 317

RESULT 7  
MTU318816  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
source  
gene  
CDS

MTU318816 637 bp DNA circular BCT 09-AUG-2002  
Mycobacterium tuberculosis partial rpoB gene for RNA polymerase  
beta subunit, isolate 1255-98.  
AJ318816  
AJ318816.1 GI:22208408  
RNA polymerase beta subunit; rpoB gene.  
Mycobacterium tuberculosis.  
Mycobacterium tuberculosis.  
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium  
tuberculosis complex.  
Herrera, L., Jimenez, M.S. and Saez, J.A.  
Molecular analysis of rifampin-resistant Mycobacterium tuberculosis  
isolated in Spain (1996-2001). Description of new alleles into rpoB  
gene and review  
Unpublished  
2 (bases 1 to 637)  
Herrera, L.  
Direct Submission  
Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro  
Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5,  
Majadahonda, Madrid. 28220, SPAIN

Location/Qualifiers  
1..637  
/organism="Mycobacterium tuberculosis"  
/isolate="1255-98"  
/db\_xref="taxon:1773"  
1..637  
/gene="rpoB"  
<1..>637  
/codon\_start=1  
/transl\_table=11  
/product="RNA polymerase beta subunit"  
/protein\_id="CAC87033.1"  
/db\_xref="GI:22208409"

/translation="VGTDEALLDIYRKLPGSPPTKESAOQLLENLFFKEKRYDLARV  
GRYKNNKGLHVGEPITSTLTEDVATIEYLRHHEGOTMTVPVGVPEVPTDD  
IDHFGNRRRTVGLIQNIRVGMRSRMRVVRMTTQDVEAITPOTLINIRPVVAI  
KEFFGTQSQFMGQNNPLSLTKRRLSALPGGLSRERAGLEVRDVHPSH"  
BASE COUNT 128 a 198 c 210 g 101 t

Query Match 80.0%; Score 166.4; DB 1; Length 637;  
Best Local Similarity 87.5%; Pred. No. 1.1e-21;  
Matches 182; Conservative 0; Mismatches 26; Indels 0; Gaps 0;  
ORIGIN  
QY 1 TCAAGGAGAGCGCTACGACCTGGCCGCTGTCGCCGATACAAAGGTCACAAAGAGCTGG 60

```

Db 101   TCAAGGAGAGCGGTACGACCTTGGCCCGCTCGGTATAGGTCAACAAGAGCTCG 160
        |||
QY 61    GCCTGAACACCAATCATCGATCACCAGCAGCAGCTGACCGAAGACGTGTCGGCCA 120
        |||
Db 161   GCGTGCATGTCGGCGAGGCCATCAGCTGCTGACGCTGACCGAAGACGTCGTGGCCA 220
        |||
QY 121   CCATCGAGTATCTGGTCCGCTGACAGAGGGCCAGGCCACGATGACCGTGCAGCGCGGG 180
        |||
Db 221   CCATCGAATATCTGTCGCTTGCAGGAGGTTCAGACCAACGATGACCGTTCGGGCGGCG 280
        |||
QY 181   TCGAGGTCCGCGTGAACACCGAGCAT 208
        |||
Db 281   TCGAGGTCCGCGTGAACACCGAGCAT 308
        |||

RESULT 8
LOCUS    MTU318821 639 bp DNA circular BCT 09-AUG-2002
DEFINITION Mycobacterium tuberculosis partial rpoB gene for RNA polymerase
            beta subunit, isolate 1071-98.
ACCESSION AJ318821
VERSION    AJ318821.1 GI:22208418
KEYWORDS   RNA polymerase beta subunit; rpoB gene.
SOURCE     Mycobacterium tuberculosis.
ORGANISM   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
            Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
            tuberculosis complex.

REFERENCE
AUTHORS    Herrera, L., Jimenez, M.S. and Saez, J.A.
TITLE       Molecular analysis of rifampin-resistant Mycobacterium tuberculosis
            isolated in Spain (1996-2001). Description of new alleles into rpoB
            gene and review
JOURNAL    Unpublished
AUTHORS     2 (bases 1 to 639)
TITLE       Direct Submission
JOURNAL    Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro
            Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5.,
            Majadahonda. Madrid. 28220, SPAIN
FEATURES   Location/Qualifiers
            source          1..639
                        /organism="Mycobacterium tuberculosis"
                        /isolate="1071-98"
                        /db_xref="taxon:1773"
            gene            1..639
            CDS              <1..>639
                        /gene="rpoB"
                        /codon_start=1
                        /transl_table=11
                        /product="RNA polymerase beta subunit"
                        /protein_id="CAC87038.1"
                        /db_xref="GI:22208419"
                        /translation="RRIDEALLDIYRKLRPGEPTKESAQTLELFFKRYDLARV
                        GRYKWNKGLHVGEDITSSTLEEDVATIEYLVRLHEGQTTMTPVGGVEPVETDD
                        IDHFGNRLRTGELIQNIRVMSRMRVRMTQDVEAITPGLTNIRPAVAAI
                        KEFFGTSQLSQFMVQNNPLSLGTLTKRRLSALGPGLSRERAGLEVRDVHPSHS"
BASE COUNT 126 a 202 c 212 g 99 t
ORIGIN
Query Match 80.0%; Score 166.4; DB 1; Length 639;
Best Local Similarity 87.5%; Pred. No. 1.le-21;
Matches 182; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 1 TCAAGGAGAGCGGTACGACCTGCGCCGCTCGGTATAGGTCAACAAGAGCTCG 60
Db 101 TCAAGGAGAGCGGTACGACCTGCGCCGCTCGGTATAGGTCAACAAGAGCTCG 60
QY 101 TCAAGGAGAGCGGTACGACCTGCGCCGCTCGGTATAGGTCAACAAGAGCTCG 160
Db 101 TCAAGGAGAGCGGTACGACCTGCGCCGCTCGGTATAGGTCAACAAGAGCTCG 160
QY 61 GCCTGAACACCAATCATCGATCACCAGCAGCAGCTGACCGAAGACGTGTCGGCCA 120
Db 161 GCGTGCATGTCGGCGAGGCCATCAGCTGCTGACGCTGACCGAAGACGTCGTGGCCA 220
QY 121 CCATCGAGTATCTGGTCCGCTGACAGAGGGCCAGGCCACGATGACCGTGCAGCGCGGG 180
Db 221 CCATCGAATATCTGTCGCTTGCAGGAGGTTCAGACCAACGATGACCGTTCGGGCGGCG 280
QY 181 TCGAGGTCCGCGTGAACACCGAGCAT 208
Db 281 TCGAGGTCCGCGTGAACACCGAGCAT 308

RESULT 9
LOCUS    I50706 970 bp DNA linear PAT 07-OCT-1997
DEFINITION Sequence 1 from patent US 5643723.
ACCESSION I50706
VERSION    I50706.1 GI:2472409
KEYWORDS   .
SOURCE     Unknown.
ORGANISM   Unclassified.
REFERENCE  1 (bases 1 to 970)
AUTHORS    Persing, D.H., Hunt, J.J., Young, K.K.Y., Felmlee, T.A., Roberts, G.D.
            and Whelan, A.Christian.
TITLE       Detection of a genetic locus encoding resistance to rifampin in
            mycobacterial cultures and in clinical specimens
JOURNAL    Patent: US 5643723-A 1 01-JUL-1997;
FEATURES   Location/Qualifiers
            source          1..970
                        /organism="unknown"
BASE COUNT 182 a 302 c 330 g 156 t
ORIGIN
Query Match 80.0%; Score 166.4; DB 6; Length 970;
Best Local Similarity 87.5%; Pred. No. 1e-21;
Matches 182; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 1 TCAAGGAGAGCGGTACGACCTGCGCCGCTCGGTATAGGTCAACAAGAGCTCG 60
Db 26 TCAAGGAGAGCGGTACGACCTGCGCCGCTCGGTATAGGTCAACAAGAGCTCG 85
QY 61 GCCTGAACACCAATCATCGATCACCAGCAGCAGCTGACCGAAGACGTCGTGGCCA 120
Db 86 GGTGTCATGTCGGCGAGCCCATCACGTCGTCGACCGAAGACGTCGTGGCCA 145
QY 121 CCATCGAGTATCTGGTCCGCTGACAGAGGGCCAGGCCACGATGACCGTGCAGCGCGGG 180
Db 146 CCATCGAATATCTGGTCCGCTTGCAGAGGTTCAGACCAACGATGACCGTTCGGGCGGCG 205
QY 181 TCGAGGTCCGCGTGAACACCGAGCAT 208
Db 206 TCGAGGTCCGCGTGAACACCGAGCAT 233

RESULT 10
LOCUS    AX111339 3534 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 2072 from Patent WO0123604.
ACCESSION AX111339
VERSION    AX111339.1 GI:13927631
KEYWORDS   .
SOURCE     Mycobacterium tuberculosis.
ORGANISM   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
            Actinomycetales; Corynebacterineae; Mycobacteriaceae;
            Mycobacterium; Mycobacterium tuberculosis complex.
REFERENCE  1 (bases 1 to 3534)
AUTHORS    Bergeron, M.G., Boissinot, M., Huletsky, A., m Nard, C., Ouellette, M.,
            Picard, F.J. and Roy, P.H.
TITLE       Highly conserved genes and their use to generate probes and primers
            for detection of microorganisms
JOURNAL    Patent: WO 0123604-A 2072 05-APR-2001;
            Infectio Diagnostica (I.D.I.) INC. (CA)
FEATURES   Location/Qualifiers

```

```

source
1. .3534
/organism="Mycobacterium tuberculosis"
/strain="Rv"
/db_xref="taxon:1773"
BASE COUNT 679 a 1081 c 1188 g 586 t
ORIGIN

Query Match 80.0%; Score 166.4; DB 6; Length 3534;
Best Local Similarity 87.5%; Pred. No. 8.9e-22;
Matches 182; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 1 TCAAGGAGAGCGCTACGACCTGGCCGCTGCGCGGATACAGGTCAACAAGAGCTGG 60
Db 902 TCAAGGAGAGCGCTACGACCTGGCCGCTGCGCGGATACAGGTCAACAAGAGCTGG 961

QY 61 GCCTGAACCAATCATCCGATCACACGACGACGCTGACCGAAGAAGACGTCGCGCCA 120
Db 962 GCCTGATGTCGCGAGCGCCATCACGTCGCGACGCTGACCGAAGAAGACGTCGCGCCA 1021

QY 121 CCATCGAGTATCTGTCGCGCTGCACGAGCGCCAGCGACGATGACCGTCGCGCGGGG 180
Db 1022 CCATCGAATATCTGTCGCGCTGCACGAGGTCAGACCAAGATGACCGTTCGCGCGGGG 1081

QY 181 TCGAGTGCGCGTGGAACCGACGACAT 208
Db 1082 TCGAGTGCGCGTGGAACCGACGACAT 1109

RESULT 11
MTU12205 3853 bp DNA linear BCT 02-MAR-2000
LOCUS
DEFINITION
Mycobacterium tuberculosis H37Rv RNA-polymerase beta subunit (rpoB)
gene, partial cds.
ACCESSION
U12205
VERSION
U12205.1 GI:515684
KEYWORDS
Mycobacterium tuberculosis.
SOURCE
Mycobacterium tuberculosis.
ORGANISM
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Corynebacterineae; Mycobacteriaceae;
Mycobacterium; Mycobacterium tuberculosis complex.
.
REFERENCE
1 (bases 1 to 3853)
Imboden, P., Trollier, R., Marchesi, F., Telenti, A., Bodmer, T.,
Cole, S., Schopfer, K. and Burkart, T.
The rpoB gene of Mycobacterium tuberculosis
Unpublished
2 (bases 1 to 3853)
Imboden, P.
Direct Submission
Submitted (11-JUL-1994) Paul Imboden, Institute for Medical
Microbiology, University of Berne, Friedbuehlstrasse 51, Berne,
3010, Switzerland
Location/Qualifiers
1. .3853
/organism="Mycobacterium tuberculosis"
/strain="H37Rv"
/db_xref="taxon:1773"
576. .>3853
/gene="rpoB"
576. .>3853
/gene="rpoB"
/codon_start=1
/transl_table=11
/product="RNA-polymerase beta subunit"
/protein_id="AAA20242.2"
/db_xref="GI:7144493"
/translation="MLEGCILADRSQSKTAASPSRQSSNNNSVPGAPNRYVFAKL
REPLEVGLLDVQDPSFEWLTGSRWRSAERGDVNPVGLLEVLIELSPIEDFSGS
MSLSFSDPRFDVDRAPYDECKDKDMTYAFLFYTAEP INNNTGEIKSOTVFMGDFPMW
TEKGTFTINGTRVSVQVYFDEITDKSTDKTLHSVKVIPSRAWLEFDYDK
RDTGVVRIDRRRROPVTVLLKALGWTSEIIVERFGESEINRSLEKDNVTGTEALLD
IYKRLRGEPTKESATLLENLFFKEKRIOLARVGRYKVNKKLGLRHVGEPITTSLT
EDDVATIEYLVRUHEGQTTMTVPVGVEVPVETDDIDHFNRRRLRTVGLIQNRVIG
MSRERVVRRMTTQDVEAITPQTILINIRPVAAAIKEFFGTQSOLFMDONPLSGILT
HKRRLSALGPGLSRRERAGLEVDRVHPSHYGRMCPTETPEGPNIGLIGLSVYRVNP
FGFTETPYRKVVDVSDGIVYLTADEEDRHVQAQNSPIDADGRVERVLYRRKAG
EVEVPSSEVDYMDVSPROMVSVATAMIPLEHDDANRALMGANMORQAVPLVRSAP
LVGTGHELRADIAIDATYSSQESGVIEVSADIYVMDHNGTRRTYRMRKPAKSNHGT
ANQCPIVDAGDRVEAGQVIADGPTDDGEMALGNLLVAIMPMEHNYEDAITLSNRL
VEEDVLTSHIEEHEIDARDLKLGAEEITRDIPNIDEVLAIDLDERGIVRIGAEVRDG
DILVGVTPKGETELTPEERLLRAIFGEKAREVRDTSKVPHGESCQVIGIRVFSRED
EDELPGAVNELVRVYVAOKRKISDQKLAGRHGNKGVICKILPVEDMPFLADGTPVDI
ILANTHGVPRMNIGQILETHLHCASHGHKVDAAKGVDPWAARLPDELLAQPNALVS
TPVDFGAQAEIAGLISCTLPNRDGDVLDADKAMLFDRGRSEPEFPYPVTVGYMTIM
KLHLHVDYDKIHKRSTGPYSMTITQPLGGKAQGGQRFGEHMECHMAQYGAAYTLOELL
TIKS"
BASE COUNT 723 a 1173 c 1293 g 664 t
ORIGIN

Query Match 80.0%; Score 166.4; DB 1; Length 3853;
Best Local Similarity 87.5%; Pred. No. 8.9e-22;
Matches 182; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 1 TCAAGGAGAGCGCTACGACCTGGCCGCTGTCGGCCGATACAGGTCAACAAGAGCTGG 60
Db 1477 TCAAGGAGAGCGCTACGACCTGGCCGCTGTCGGCTATAGGTCAACAAGAGCTCG 1536

QY 61 GCCTGAACCAATCATCCGATCACACGACGACGCTGACCGAAGAAGACGTCGCGCCA 120
Db 1537 GCCTGCATGTCGCGGAGCCCATACGTCGCGCTGACCGAAGAAGACGTCGCGCCA 1596

QY 121 CCATCGAGTATCTGTCGCGCTGCACGAGCGCCAGCGACGATGACCGTCGCGGGGGG 180
Db 1597 CCATCGAATATCTGTCGCGCTGCACGAGGTCAGACCAAGATGACCGTTCGCGGGGGG 1656

QY 181 TCGAGTGCGCGTGGAACCGACGACAT 208
Db 1657 TCGAGTGCGCGTGGAACCGACGACAT 1684

RESULT 12
MSGRPOB 5084 bp DNA linear BCT 13-SEP-1994
LOCUS
DEFINITION
Mycobacterium tuberculosis RNA polymerase beta subunit (rpoB)
gene, complete cds and RNA polymerase beta'-subunit rpoC gene,
partial cds.
L27989
L27989.1 GI:468333
ACCESSION
RNA polymerase beta-subunit; rpoB gene.
KEYWORDS
Mycobacterium tuberculosis (strain Rv) DNA.
SOURCE
Mycobacterium tuberculosis
ORGANISM
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
tuberculosis complex.
1 (bases 1 to 5084)
Miller, L.P., Crawford, J.T. and Shinnick, T.M.
The rpoB gene of Mycobacterium tuberculosis
Antimicrob. Agents Chemother. 38 (4), 803-811 (1994)
94304130
PUBMED
8031050
Location/Qualifiers
1. .5084
/organism="Mycobacterium tuberculosis"
/strain="Rv"
/db_xref="taxon:1773"
1065. .4598
/gene="rpoB"
1065. .4598
/gene="rpoB"
/codon_start=1
/evidence-experimental
/transl_table=11
/product="RNA polymerase beta-subunit"
/protein_id="AAA21416.1"
/db_xref="GI:468334"
/translation="MLEGCILADRSQSKTAASPSRQSSNNNSVPGAPNRYVFAKL

```

MSLSEFSDPDDVDKAPVDECKDKDMYAAFLVTAEFNNNTGEIKSOTVFMGDFPM  
TEKGFTIINGTERRVVSQVLSPGVFDETDIDKSTKLHSVKVIPSPGAWLEFDVDK  
RTDVGIDDKRQRPVILLKALGMTSEQIVRFGFSEIMRSTLEKDNVTGTDEALLD  
IYKLRPGDEPTKESAQTLLENLFFKEKRYDLARVGRYKVKKKLGHVGEPTTSSTLT  
EEDVATIEYVRLHHEGQTTMTVPGVEVPVETDDIDHFGNRLRTVGELEONQIRVG  
MSRMERVRERTMTQDVEAITPOTLINIRPVAAIKEFEFGTSLQSFQDQNNPLSGLT  
HKRRLSALGPGGLSRERAGLEVDRVHPSHYGRMCPIETPEGNIGLIGLSLYARVNP  
EGFLETYPKYVDGVSDVSDIYLTADDEDRHVVAOANSPIDAGDRFVPRVLRKAG  
EVEYFSSDGLRAAIDAATSSQESGVIEESADYITVMHNGTTRRTYMRKFAHSNHGTC  
LVGTGMELRAAIDAGVVAEESGVIEESADYITVMHNGTTRRTYMRKFAHSNHGTC  
ANOCPIVDAGRAAGVQIADGCTDDGEMALGNLLVAIMPWECHNTEDAILLSNRL  
VEEDVLTSHIEHEIDARDTKLGABEITRDIPNISDEVLADLDERGIVRIGAEVRDG  
DILGVKVTPEKGETLTPERLLRAIFGEKAREVRDTSKLVPHGESGKVIIGRVSRED  
EDELPAVGNELVRVYVAQKRLSDGDKLAGRHNGKVGKILPVEDMFPDLAGDTPVDI  
ILNTHGVPRMNIQOILETHLGCWCAHSGWKVDAAGVDPDMAARUPDELELEAHANVDS  
TPVFGAQAELQGLLSCTLPNRDGDVLVDADKAMLEDFGRSGEPFPYPTVGYMYIM  
KLHLVDVDRKHARSTGPYSMITQOPLGGKAQFGQRFGEWECWMAQVGAAYTLQELL  
TKSDDTYGRVRYEAIKNGENIPEPGIPESFKVLLKELQSLCLNVEVLSDDGAIEL  
REGDEDLERAANLGINLSRNESASFEDLA"

641. 5084  
/gene="rpoc"  
4641. .>5084  
/gene="rpoc"  
/codon\_start=1  
/transl\_table=11  
/product="RNA polymerase beta'-subunit"  
/protein\_id="AA21417.1"  
/db\_xref="GI:537608"  
/translation="MLDVNFDELRIGLATAEDIRQWSYGEVKKPETINRYTLKPEK  
GLFCEKIFGPTDMCYCGYKRVFKGIIICERCVEVTRAKVRERMGHIELAAPVT  
HIWYFKGVSRLGYLLDLPKOLEKIIFYAAVYVITSVDEENRHNL"

BASE COUNT 969 a 1534 c 1691 g 890 t

ORIGIN

Query Match 80.0%; Score 166.4; DB 1; Length 5084;  
Best Local Similarity 87.5%; Pred. No. 8.5e-22;  
Matches 182; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 1 TCAGGAGAACGCGTACGACCTGGCCCGTGTGCGCGGATACAGGTCAACAGAGCGTGG 60  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
Db 1966 TCAAGGAGAACGCGTACGACCTGGCCCGGTCGTCGATTAAGGTCAACAGAGCGTGG 2025  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

QY 61 GCGTCAACACCAATCATCGATCACACGACGACCTGACCCAGAACAGCGTCGTCGCCA 120  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 2026 GCGTCAATCGCGGAGCCATCATCGTCTGACGCTGACCCAGAACAGCGTCGTCGCCA 2085  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

QY 121 CCATCAGATATCTGTCGCCCTGACGAGGCGGACGACGATGACCGTGGCGGCGGG 180  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 2086 CCATCAATATCTGTCGCCCTGACGAGGCTGACACGACGATGACCGTTCGCGGCGGG 2145  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

QY 181 TCGAGTGCCTGGGAAACGACGACAT 208  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 2146 TCGAGTGCCTGGGAAACGACGACAT 2173  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

RESULT 13  
AE006964  
LOCUS  
DEFINITION  
Mycobacterium tuberculosis CDC1551, section 50 of 280 of the complete genome.  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS

AE006964  
Mycobacterium tuberculosis CDC1551, section 50 of 280 of the complete genome.  
AE006964 AE000516  
AE006964.1 GI:13880217  
Mycobacterium tuberculosis CDC1551.  
Mycobacterium tuberculosis CDC1551.  
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
Actinomycetales; Corynebacteriaceae; Mycobacteriaceae;  
Mycobacterium tuberculosis complex.  
1 (bases 1 to 19352)  
Fleischmann,R.D., Alland,D., Eisen,J.A., Carpenter,L., White,O., Peterson,J., DeBoy,R., Dodson,R., Gwinn,M., Haft,D., Hickey,E., Kolonay,J.F., Nelson,W.C., Umayam,L.A., Ermolaeva,M.,

Salzberg,S.L., Delcher,A., Utterback,T., Weidman,J., Khouiri,H., Gill,J., Mikula,A. and Bishai,W.  
Whole genome comparison of Mycobacterium tuberculosis clinical and laboratory strains  
Unpublished  
2 (bases 1 to 19352)  
Fleischmann,R.D., Alland,D., Eisen,J.A., Carpenter,L., White,O., Peterson,J., DeBoy,R., Dodson,R., Gwinn,M., Haft,D., Hickey,E., Kolonay,J.F., Nelson,W.C., Umayam,L.A., Ermolaeva,M., Hickey,E., Salzberg,S.L., Delcher,A., Utterback,T., Weidman,J., Khouiri,H., Gill,J., Mikula,A. and Bishai,W.  
Direct Submission  
Submitted (25-APR-2001) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA  
Location/Qualifiers  
1. 19352  
/organism="Mycobacterium tuberculosis CDC1551"  
/strain="CDC1551"  
/db\_xref="taxon:83331"  
/note="clinical strain"  
163. .3699  
/gene="MT0695"  
163. 3699  
/gene="MT0695"  
/note="similar to GB:L27989 GB:L05910 GB:U12205 SP:P47766 PID:149992; identified by sequence similarity; putative"  
/codon\_start=1  
/transl\_table=11  
/product="DNA directed RNA polymerase, beta subunit"  
/protein\_id="AAK44921.1"  
/db\_xref="GI:13880218"  
/translation="MLEGCIADRSQKTAASPSRPOSSNSVPGAPNRVVSFAKL  
REPLEVPLGLDVOTDGFSPRMESAAERGDVNPVGLLEEVLYELSPIDFSGS  
MSLSEFSDPDDVDKAPVDECKDKDMYAAFLVTAEFNNNTGEIKSOTVFMGDFPM  
TEKGFTIINGTERRVVSQVLSPGVFDETDIDKSTKLHSVKVIPSPGAWLEFDVDK  
RTDVGIDDKRQRPVILLKALGMTSEQIVRFGFSEIMRSTLEKDNVTGTDEALLD  
IYKLRPGDEPTKESAQTLLENLFFKEKRYDLARVGRYKVKKKLGHVGEPTTSSTLT  
EEDVATIEYVRLHHEGQTTMTVPGVEVPVETDDIDHFGNRLRTVGELEONQIRVG  
MSRMERVRERTMTQDVEAITPOTLINIRPVAAIKEFEFGTSLQSFQDQNNPLSGLT  
HKRRLSALGPGGLSRERAGLEVDRVHPSHYGRMCPIETPEGNIGLIGLSLYARVNP  
EGFLETYPKYVDGVSDVSDIYLTADDEDRHVVAOANSPIDAGDRFVPRVLRKAG  
EVEYFSSDGLRAAIDAATSSQESGVIEESADYITVMHNGTTRRTYMRKFAHSNHGTC  
LVGTGMELRAAIDAGVVAEESGVIEESADYITVMHNGTTRRTYMRKFAHSNHGTC  
ANOCPIVDAGRAAGVQIADGCTDDGEMALGNLLVAIMPWECHNTEDAILLSNRL  
VEEDVLTSHIEHEIDARDTKLGABEITRDIPNISDEVLADLDERGIVRIGAEVRDG  
DILGVKVTPEKGETLTPERLLRAIFGEKAREVRDTSKLVPHGESGKVIIGRVSRED  
EDELPAVGNELVRVYVAQKRLSDGDKLAGRHNGKVGKILPVEDMFPDLAGDTPVDI  
ILNTHGVPRMNIQOILETHLGCWCAHSGWKVDAAGVDPDMAARUPDELELEAHANVDS  
TPVFGAQAELQGLLSCTLPNRDGDVLVDADKAMLEDFGRSGEPFPYPTVGYMYIM  
KLHLVDVDRKHARSTGPYSMITQOPLGGKAQFGQRFGEWECWMAQVGAAYTLQELL  
REGDEDLERAANLGINLSRNESASFEDLA"

3744. .7694  
/gene="MT0696"  
3744. 7694  
/gene="MT0696"  
/note="similar to SP:P37871; identified by sequence similarity; putative"  
/codon\_start=1  
/transl\_table=11  
/product="DNA-directed RNA polymerase, beta-prime subunit"  
/protein\_id="AAK44922.1"  
/db\_xref="GI:13880219"  
/translation="MLDVNFDELRIGLATAEDIRQWSYGEVKKPETINRYTLKPEK  
GLFCEKIFGPTDMCYCGYKRVFKGIIICERCVEVTRAKVRERMGHIELAAPVT  
HIWYFKGVSRLGYLLDLPKOLEKIIFYAAVYVITSVDEENRHNLSTLEAEAVRER  
AVEDQKQGLEARQKLEADLAELAEAGAKADARRKRDGGERMRQINDRAQRELD  
LEDIWTFTKLPKQIYVDENLYRELVDYRGYFTGAMGASIQIKLIENFIDAEAS  
LRDVIKNGKQKRLKRLKLVAAFOQSGNSPMGVLDAPVPIPELLRPMVLOGDGR  
FATSDNLNLYRRVINNRNLKRLIDLGAPEIIVNNEKRLQESVDALFQNGRGPVY  
GPNRPLKSLDLKKGQGRFQNLGKRVDSGRSVIVGVQPLKHGCGLPKLMAL  
LKPFVPMKRLVDLNRHQAQISAKRVWRQRPQMDVLEEVIAEHVPLNLRHLRLG  
IQAFEPMLVEGKAIQLHPLVCEAFNAFDGDMQVHLPLSAEQAQAEARILMLSSNNIL





REFERENCE  
AUTHORS

1 (bases 1 to 19770)  
Cole, S. T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C.,  
Harris, D., Gordon, S. V., Eigemeier, K., Gas, S., Barry III, C. E.,  
Tekai, F., Badcock, K., Basham, D., Brown, D., Chillingworth, T.,  
Connor, R., Davies, K., Devlin, K., Feltwell, T., Gentles, S.,  
Hamlin, N., Holroyd, S., Hornsby, T., Jagels, K., Krogh, A., McLean, J.,  
Moule, S., Murphy, L., Oliver, S., Osborne, J., Quail, M. A.,  
Rajandream, M. A., Rogers, J., Rutter, S., Seeger, K., Skelton, S.,  
Squares, S., Squires, R., Sulston, J. E., Taylor, K., Whitehead, S. and  
Barrell, B. G.  
Deciphering the biology of Mycobacterium tuberculosis from the  
complete genome sequence  
Nature 393 (6685), 537-544 (1998)  
98295987  
9634230  
2 (bases 1 to 19770)  
Parkhill, J.  
Direct Submission  
Submitted (11-JUN-1998) Submitted on behalf of the Mycobacterium  
tuberculosis sequencing and mapping teams. Sanger Centre, Wellcome  
Trust Genome Campus, Hinxton, Cambridge CB10 1SA Unite de Genetique  
Moleculaire Bacterienne, Institut Pasteur, 28 rue du Docteur Roux,  
75724 Paris Cedex 15, France E-mail: parkhill@sanger.ac.uk  
On Jun 27, 1998 this sequence version replaced gi:2143285.  
Notes:  
Details of M. tuberculosis sequencing at the Sanger Centre are  
available on the World Wide Web.  
(URL, <http://www.sanger.ac.uk/Projects/M.tuberculosis/>) CDS have  
been renumbered from the original cosmid submissions but the old  
gene designations are in brackets after the new gene numbers.  
Gene prediction was based on a Hidden Markov Model of TB genes  
implemented in TBparse (Krogh) supplemented with visual inspection  
of positional base preference in codons, especially where there is  
an increase in the observed/expected third position G + C.  
CAUTION: In some cases we may not have predicted the correct  
initiation codon. Where possible we choose an initiation codon  
(atg, gtg, or ttg) which is preceded by an upstream ribosome  
binding site sequence (optimally 5-13bp before the initiation  
codon). If this cannot be identified we choose the most upstream  
initiation codon.

## COMMENT

FEATURES  
source

Location/Qualifiers  
1. .19770  
/organism="Mycobacterium tuberculosis H37Rv"  
/strain="H37Rv"  
/db\_xref="taxon:83332"  
/clone="1376"  
57. .61  
/note="aggA, possible rbs upstream of Rv0654"  
68. .1573  
/gene="Rv0654"  
68. .1573  
/gene="Rv0654"  
/note="Rv0654", (MTCI376.22), len: 501. unknown, FASTA  
score: Q53353 LIGNOSTILBENE-ALPHA, BETA-DIOXYGENASE (485  
aa) opt: 280 z-score: 330.1 E(): 2.3e-11, (28.5% identity  
in 523 aa overlap). Also similar to M. tuberculosis protein  
MTCY212.07c (29.5% identity in 522 aa overlap)."  
/codon\_start=1  
/transl\_table=1  
/product="hypothetical protein Rv0654"  
/protein\_id="CAB09380.1"  
/db\_xref="GI:2143307"  
/db\_xref="SPTREMBL:O06785"  
/translation="MTTAAAESQNPYLEGLFAPVSTVATDLPVTGRIPHELDGRY  
LNGPNPAEVDPAHYHFTGDGMVGVNLRGDKARWYRNVRVTPCAALGEFISA  
RHPRTGIIEGPNVNLTHAGTALVLAAGVNYEYDELDTGPGDFDGLTHGGY  
APDQRTGIELHVSYSFARGHRYQVSVIGTDHARRVDIEVAGSPMHHSFLTND  
YVYIDLPTFDPQVPVSPASVPRWLQRPRLVQSVLGRVTRIPDIAALGNRMVQSHD  
RLPYAMNPSYPARVGVMPREGNEDRWFIPEPCVYHPLNAYSRCNGAEVILVDVY  
RYSRMFDRRCGGDSRPSLDRTWTLNLTAGVTAECRDDRAQEPRIENILVGGPHR  
FAYTVGIEGFLVGGAALSTPLYKQDCVTSVTSVLSLDPDLLEGEMVFVFNPSARAE  
DGIILMGVGNHGRGREGQLLLDADLTLESIAIVHLPQVPMVGFHGNWAPT"  
1585. .2664

## gene

## CDS

/gene="Rv0655"  
1585. .2664  
/gene="Rv0655"  
/note="Rv0655", (MTCI376.21), len: 359. abc transporter,  
FASTA score: YRBF\_ECOLI\_P45393 hypothetical abc  
transporter atp-binding (269 aa) opt: 644 z-score: 721.8  
E(): 3.4e-33 (38.5% identity in 244 aa overlap); contains  
PS00017 ATP/GTP-binding site motif A, PS00211 ABC  
transporters family signature, highly similar to M. leprae  
MKL\_MYCLE\_P30769 possible ribonucleotide transport atp-  
(347 aa) opt: 2021 z-score: 2244.4 E(): 0, (92.2% identity  
in 335 aa overlap). Also similar to many other M.  
tuberculosis ABC transporters eg. MTCY253.24 (33.6%  
identity in 241 aa overlap)."  
/codon\_start=1  
/transl\_table=1  
/product="hypothetical protein Rv0655"  
/protein\_id="CAB09379.1"  
/db\_xref="GI:2143306"  
/db\_xref="SWISS-PROT:O06784"  
/translation="MRYSDSYHTGRMQPRASTGEFPMGVSIEVNGLTKSFGSSRIWE  
DVTLTIPAGEVSLVLLPGSGTKSVFLKSLIGLLRPERGSIIDGTDITIECSAKELYEI  
RTLGVLFQDQALFGSMNLXNTAFPLREHTKKKESEIRDIYMEKLAIVGLGGEKFF  
PGEISGGMRKAGLARALVLDPQIILCDPDSGLDPVATYLSQLINDINAOIATIL  
IVYTHNARTVPDNNMGLFRKHLVMEFGFREVLLTSDFVVRQFLNGRIPGIMSEE  
KDEATMAEEQALLDAGHHAGGVEETGEPVPOISATPGMPERKAVARQARVREMLHTL  
PKKAAAILDDLEGTHKYAVHEIQG"  
1762. .1785  
/gene="Rv0655"  
/note="PS00017 ATP/GTP-binding site motif A"  
2074. .2118  
/gene="Rv0655"  
/note="PS00211 ABC transporters family signature"  
complement(3052. .3435)  
/gene="Rv0656c"  
complement(3052. .3435)  
/gene="Rv0656c"  
/note="Rv0656c", (MTCI376.20, unknown), len: 127 aa"  
/codon\_start=1  
/transl\_table=1  
/product="hypothetical protein Rv0656c"  
/protein\_id="CAB09378.1"  
/db\_xref="GI:2143305"  
/db\_xref="SPTREMBL:O06783"  
/translation="MAATTTGTHRGLELRAAQRAGVSGCEPQRAEFCRSARNADFFDQ  
MSRMFGDVPDPVPKSVYMRIDISQAHLRAGAVGALSVVVDLLICDTAAARGLVWLH  
DDADYELAERHLPDIRRVVSADD"  
complement(3530. .3685)  
/gene="Rv0657c"  
complement(3530. .3685)  
/note="Rv0657c", (MTCI376.19), unknown, len: 51 aa; similar  
to several other M. tuberculosis hypothetical proteins eg.  
YW08\_MYCTU\_Q10848 hypothetical 8.9 kd protein cy39.08c (80  
aa) fasta scores; opt: 107 z-score: 182.3 E(): 0.0038,  
45.8% identity in 48 aa overlap. Also similar to MTCY48\_5  
and AL020958|SC4H8\_7 Streptomyces coelicolor cosmid 4H8 (66  
aa), 41.0% identity in 39 aa overlap."  
/codon\_start=1  
/transl\_table=1  
/product="hypothetical protein Rv0657c"  
/protein\_id="CAB09377.1"  
/db\_xref="GI:2143304"  
/db\_xref="SPTREMBL:O06782"  
/translation="MSVTQIDLDELADVMRIAIVHTKKKAVNLAMRDYVERFRRIE  
ALARSRE"  
complement(3691. .3696)  
/note="possible RBS upstream of Rv0657c"  
complement(3761. .4477)  
/gene="Rv0658c"  
complement(3761. .4477)  
/gene="Rv0658c"  
/note="Rv0658c", (MTCI376.18), len: 238, unknown,





Db 1248 TCGAGTTGCGCGTCGAGTCGCACGACAT 1275

Search completed: November 13, 2002, 01:25:56  
Job time : 662.723 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 12, 2002, 16:59:42 ; Search time 1141.13 Seconds  
(without alignments)  
3037.202 Million cell updates/sec

Title: US-09-697-123B-7  
Perfect score: 214  
Sequence: 1 tcaagggaagcgtacgac.....ccggtcaggtcagcagat 214

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues  
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estum:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vrt:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rod:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES					
Result No.	Score	Query Match	Length	ID	Description
c 1	139.2	65.0	1282	9 AI770311	AI770311 42 Mycoba
2	47.4	22.1	1850	11 AY107051	AY107051 Zea mays
c 3	45	21.0	569	17 BH897054	BH897054 3526_1_6
4	44.6	20.8	665	13 BM601475	BM601475 170006590
5	44.6	20.8	687	13 BM590640	BM590640 170006873
6	43.8	20.5	275	14 R86542	R86542 RABEST103T

7	43.6	20.4	371	10	AV438805
8	43.2	20.2	461	14	C73000
9	43	20.1	234	14	R86588
10	43	20.1	235	14	R86566
11	43	20.1	331	14	R86510
12	43	20.1	878	12	BF264852
13	42.6	19.9	618	13	BJ270031
14	42.6	19.9	623	13	BJ270031
15	42.4	19.8	479	10	BE361556
16	42.2	19.7	279	14	R86528
17	42.2	19.7	710	12	BG321153
18	42	19.6	2598	11	AY103647
19	41.6	19.4	1856	11	AY109418
20	41.4	19.3	416	12	BF484291
21	41.4	19.3	714	14	BQ842534
22	40.8	19.1	1006	12	BG106122
23	40.8	19.1	1750	11	BC029105
24	40.6	19.0	434	13	BI578713
c 25	40	18.7	257	17	BH804652
26	39.8	18.6	407	10	BE406021
c 27	39.8	18.6	525	10	AV929492
28	39.8	18.6	550	13	BJ208458
29	39.8	18.6	558	12	BE754521
30	39.8	18.6	579	13	BJ208491
31	39.8	18.6	585	17	BH806894
32	39.8	18.6	589	13	BJ244734
33	39.8	18.6	613	13	BJ260431
34	39.8	18.6	624	13	BJ248032
35	39.8	18.6	633	13	BJ317529
36	39.8	18.6	645	10	AW056150
37	39.8	18.6	645	14	BQ241492
38	39.8	18.6	672	13	BJ244763
39	39.8	18.6	680	14	BQ820533
40	39.8	18.6	792	12	BF627944
41	39.8	18.6	840	10	AW448084
42	39.8	18.6	840	10	BQ606019
43	39.8	18.6	925	17	CNS0091P
44	39.6	18.5	417	10	BE363166
45	39.6	18.5	500	10	BE360701

ALIGNMENTS

RESULT 1  
AI770311/c 1282 bp mRNA linear EST 24-JAN-2000  
LOCUS AI770311  
DEFINITION 42 Mycobacterium anaerobic stationary phase library Mycobacterium smegmatis cDNA, mRNA sequence.  
ACCESSION AI770311 GI:6742680  
VERSION AI770311.1  
KEYWORDS EST.  
SOURCE Mycobacterium smegmatis.  
ORGANISM Mycobacterium smegmatis.  
REFERENCE 1 (bases 1 to 1282)  
AUTHORS Murugasu-Oei, B., Tay, A. and Dick, T.  
TITLE Upregulation of stress response genes and ABC transporters in anaerobic stationary-phase Mycobacterium smegmatis  
JOURNAL Mol. Gen. Genet. 262 (4-5), 677-682 (1999)  
MEDLINE 20092472  
COMMENT Contact: Murugasu-Oei, B.  
Mycobacterium Laboratory  
Institute of Molecular and Cell Biology  
30 Medical Drive, Singapore 117609, Republic of Singapore  
Tel: 65 874 3011  
Fax: 65 779 1117  
Email: mcbome@imcb.nus.edu.sg  
Insert Length: 1282 Std Error: 0.00  
Seq primer: T3 Forward; T7 Backward.

FEATURES  
source location/Qualifiers  
1. .1282

```

/organism="Mycobacterium smegmatis"
/strain="mc2155"
/db_xref="taxon:1772"
/clone_lib="Mycobacterium anaerobic stationary phase
library"
/lab_host="E. coli XL1-Blue MRF"
/notes="Vector: Lambda ZAP II; Bacilli were disrupted using
a Mini bead beater (Biospec). RNA was isolated using the
RNeasy protocol (Qiagen). Purified RNA was subjected to 2
rounds of digestion with RNase-free DNase I (Promega).
DNase I was heat-inactivated at 75oC for 5 min. and
removed by using RNeasy columns followed by phenol
extraction and ethanol precipitation. The RNA
preparations were confirmed to be free of genomic DNA
contamination by carrying out PCR and RT-PCR using the
Access kit (Promega) and primers specific for the
histone-like protein gene hlp (Lee et al., 1998). cDNA
was synthesized using random hexamer primers (Promega)
and Stratagene's cDNA synthesis kit. cDNA fragments were
ligated into lambda ZAP II vector and packaged in vitro
using Stratagene reagents."
BASE COUNT      230 a 395 c 405 g 252 t
ORIGIN

Query Match      65.0%; Score 139.2; DB 9; Length 1282;
Best Local Similarity 81.8%; Pred. No. 7.4e-21;
Matches 175; Conservative 0; Mismatches 33; Indels 6; Gaps 1;

QY 1 TCAAGGAGAGCGCTACGACCTGGCCGCGTGGCGGTACAAAGTGAACAAGAGCTGG 60
Db 236 TCAAGGAGAGCGCTACGACCTGGCCGCGTGGCGGTACAAAGTGAACAAGAGCTGG 177

QY 61 GTCTTGGCGGTGCAACCCGCGTCTCGTGACTGCCACACGCTCACCGGAGGAGACGTGC 120
Db 176 GCCTGAACGGGCAAGCC-----GATCACCAGCTCGACGCTGACCGAAGAGGACGTGC 123

QY 121 TCCCCACCATCGGTACCTGGTGGCGCTGCACGAGGGCCAGACACGATGACCCGCCCGC 180
Db 122 TCCGACCATCGAGTACCTGGTGGCGTCTGCACGAGGGTTCAGACTTCGATGACCGTCCCG 63

QY 181 GCGGCTCGAGGTCGCGGTCGAGGTCGACGACAT 214
Db 62 GTGGCGTCGAGGTTCCCGTCGAGGTCGACGACAT 29

RESULT 2
AY107051
LOCUS      AY107051      1850 bp      mRNA      linear      HTC 25-MAY-2002
DEFINITION Zea mays PCO135033 mRNA sequence.
ACCESSION  AY107051
VERSION     AY107051.1  GI:21210129
KEYWORDS   HTC.
SOURCE     Zea mays.
ORGANISM   Zea mays.

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.

AUTHORS    Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,
            Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
TITLE      Maize Mapping Project/DuPont Consensus Sequences for Design of
            Overgo Probes
JOURNAL    Unpublished (2002)
REFERENCE  2 (bases 1 to 1850)
AUTHORS    Coe,E.C.
TITLE      Direct Submission
JOURNAL    Submitted (25-APR-2002) Maize Mapping Project, University of
            Missouri, Columbia, MO 65211, USA
            Location/Qualifiers
            1..1850
            /organism="Zea mays"
            /db_xref="MaizeDB:637903"
            /db_xref="taxon:4577"

```

```

/clone="PCO135033"
/clone_lib="Maize Mapping Project/DuPont Consensus
Library"
/notes="this sequence is part of a project of EST
assemblies resulting from the application of public
contigs to seed DuPont contigs; this resource was
assembled by DuPont as part of a collaboration for the
overgo addressing of BACs in conjunction with the Maize
Mapping Project"
BASE COUNT      331 a 625 c 583 g 311 t
ORIGIN

Query Match      22.1%; Score 47.4; DB 11; Length 1850;
Best Local Similarity 52.2%; Pred. No. 0.96;
Matches 105; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

QY 12 CGCTACGACCTGGCCGCGTGGCGGTGAAAGTGAACAAGAGCTGGGCTTGGCGGT 71
Db 260 CGGTCTACCAAGAGTGGCGCGCGGTACGGGCCCATCATGACGGTGTGGCTGGGCACG 319

QY 72 GCCAACCCGCTCTGGTGACTGCCACACGCTCACCGGAGGAGACGCTCGTCCGCCACCATC 131
Db 320 TCGCCACGGTGGTGTGTCTCCACGCTCGGAGCTGGCCRAAGGAGGTGCTCAAGACCCACGAC 379

QY 132 GGTACCTGGTGGCGCTCCACGAGGCGACACGACGATGACCCGCCCGCGGCTCGAG 191
Db 380 CAGCAGCTCGCGACGCGTGGCGGCGCTCCACCGAGAGCTTCAGCCGAGGCGGCAG 439

QY 192 GTCCCGGTGCGAGGTCGACGAC 212
Db 440 GACCTCATCTGGCGGACTAC 460

RESULT 3
BH897054/c
LOCUS      BH897054      569 bp      DNA      linear      GSS 14-AUG-2002
DEFINITION 3526_1_6_1_E11.1EL_y_1 3526 - RescueMu Grid K Zea mays genomic, DNA
            sequence.
ACCESSION  BH897054
VERSION     BH897054.1  GI:22232454
KEYWORDS   GSS.
SOURCE     Zea mays.
ORGANISM   Zea mays.

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.

AUTHORS    Walbot,V.
TITLE      Maize genomic sequences found using engineered RescueMu transposon
JOURNAL    Unpublished (2001)
COMMENT    Contact: Walbot V
            Department of Biological Sciences
            Stanford University
            855 California Ave, Palo Alto, CA 94304, USA
            Tel: 650 723 2227
            Fax: 650 725 8221
            Email: walbot@stanford.edu
            Very probable ligation site of ends cut by single endonuclease.
            Reverse complemented post-ligation sequence from source sequence.
            Plate: 3526_1_6_1 row: 6
            Class: transposon-tagged.
            Location/Qualifiers
            1..569
            /organism="Zea mays"
            /cultivar="mixed background W23/A188/B73"
            /db_xref="taxon:4577"
            /clone_lib="3526 - RescueMu Grid K"
            /tissue_type="leaf"
            /dev_stage="adult"
            /lab_host="DH10B"
            /note="Organ: leaf; Vector: RescueMu (engineered from
            pBlueScript backbone); Site_1: BamHI; Site_2: BglII;
            RescueMu is a 4.9 kb, modified maize Mu transposon

```



Matches	104:	Conservative	0:	Mismatches	99:	Indels	0:	Gaps	0:
QY	8	GAAGCGCTACGACCTGGCGCCCGTGGCGGTACAAGGTGAACAAAGCTCGGTCTTTGG	67						
Db	255	GCAGCGCGACCTGGTGGCGCTGATGGCGCGCAACTCCGACGACGTGGCGCGGTGCTGCT	314						
QY	68	CGTGCCCAACCGGCTCTGGTACTGCCACACACCTCACCAGGAGACGTGCTGCGCAC	127						
Db	315	CGGTTGCTTCCTGGCGCGGTGTAAACGGTCAGCACGCTCGATCCGTCGTTTCGGCTGGAGGA	374						
QY	128	CATCGGGTACCTGGTGCOCCTCTGCAGGAGGCGAGACACGATGACCGCCCGCGCGGCT	187						
Db	375	GGTGAGCACCTGCTGCGCTGACCGCGCCGCAACGTGATAGCGGATCGGACGCGCT	434						
QY	188	CGAGTCCCGGTCGAGGTCGAGC	210						
Db	435	GCCGCTGCTACGAAGCAGCCG	457						
RESULT	6								
LOCUS	R86542								
DEFINITION	R86542								
ACCESSION	R86542								
VERSION	R86542.1								
KEYWORDS	EST								
SOURCE	rabbit								
ORGANISM	Oryctolagus cuniculus								
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;								
AUTHORS	Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.								
TITLE	1 (bases 1 to 275)								
JOURNAL	Sakai,D., Tong,H.-S. and Minkin,C.								
MEDLINE	Osteoclast Molecular Phenotyping by Random cDNA Sequencing								
COMMENT	Bone 17 (2), 111-119 (1995)								
	96021365								
	Other-ESTs: RABEST038T, RABEST045T, RABEST084T, RABEST140T,								
	RABEST163T, RABEST202T, RABEST229T								
	Contact: Sakai D								
	Basic Sciences								
	University of Southern California								
	USC School of Dentistry, 925 West 34th Street, DEN-4220, Los								
	Angeles, CA 90089-0641								
	Tel: 2137405563								
	Fax: 2137407560								
	Email: sakai@molbio.usc.edu								
	Seq primer: T7 promoter.								
FEATURES	Location/Qualifiers								
source	1..275								
	/organism="Oryctolagus cuniculus"								
	/strain="New Zealand White"								
	/db_xref="taxon:9986"								
	/clone="PRABOC103"								
	/lab_host="E. coli DH12S"								
	/notes="Vector: pSPORT1; Site_1: Salt; Site_2: NotI; Poly(A								
	) + RNA was purified from a 97% pure population of								
	osteoclasts prepared from the long bones of 10 day old								
	rabbits. First strand cDNA was synthesized by priming								
	with an oligo(dT)-NotI anchor-primer and second strand								
	cDNA was synthesized by replacement synthesis as described								
	by Gubler and Hoffman (Gene 25:283, 1983). Following the								
	addition of SalI adapters and NotI digestion, the cDNA was								
	cloned between the SalI (50) and NotI (30) sites of the								
	pSPORT1 (BRL) plasmid vector."								
BASE COUNT	67 a	76 c	92 g	40 t					
ORIGIN									
Query Match	20.5%;	Score 43.8;	DB 14;	Length 275;					
Best Local Similarity	53.1%;	Pred. NO. 4.6;							
Matches	93; Conservative	0; Mismatches	82; Indels	0; Gaps	0;				
QY	5	GGAGAAGCGCTACGACTGGCCCGCGTGGCGCGGTACAGGTGAACAAAGCTGGGTCT	64						

```

SOURCE
ORGANISM      Oryza sativa (japonica cultivar-group).
               Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhacordeae; Oryzoideae; Oryza.
1 (bases 1 to 461)
Sasaki, T. and Yamamoto, K.
Rice cDNA from panicle at flowering stage
Unpublished (1996)
Contact: Takuji Sasaki
National Institute of Agrobiological Resources
Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki
305-8602, Japan
Tel: 81-298-38-7441
Fax: 81-298-38-7468
Email: tsasaki@abr.affrc.go.jp, URL: http://rgp.dna.affrc.go.jp/
PROJECT = RGP.
Location/Qualifiers
1..461
/organism="Oryza sativa (japonica cultivar-group)"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="E2649_1A"
/clone_lib="Rice panicle at flowering stage"
/dev_stage="flowering stage"
/note="Organ: panicle; Rice cDNA from panicle at flowering
stage"
BASE COUNT    110 a 114 c 183 g 53 t 1 others
ORIGIN
Query Match    20.2%; Score 43.2; DB 14; Length 461;
Best Local Similarity 56.2%; Pred. No. 6.7;
Matches 81; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 68 CGGTCCCAACCGGCTCGTACTGCCACACGCTACCGAGGAAGAGTCGTGCGCAC 127
    || || || || || || || || || || || || || || || || || || ||
Db 21 CGAGCGAACACGCGGCTGGTGTGCGGCGGTGGAGGAAGAGCGCGAACAC 80
    || || || || || || || || || || || || || || || || || || ||
QY 128 CATCCGGTACCTGTGTCGCGCTGCACGAGGCGCAGACACGATACCGCGCGCGCT 187
    || || || || || || || || || || || || || || || || || || ||
Db 81 GATCAAGTCGCGGAGCGGAGAGGAGATCGGACACACCGCCACCATCATCGAAGA 140

QY 188 CGAGTCCCGGTGAGGTGCGACGA 211
    || || || || || || || || || || || || || || || || || || ||
Db 141 GGAGGAGGTGTCGACGACGACGA 164

RESULT 9
R86588
LOCUS          R86588                234 bp      mRNA      linear      EST 17-AUG-1995
DEFINITION    RABEST163T Rabbit Osteoclast, Dennis Sakai Oryctolagus cuniculus
               cDNA clone PRABOC163 5' similar to cofillin, mRNA sequence.
ACCESSION     R86588
VERSION       R86588.1 GI:947242
KEYWORDS      EST.
SOURCE        rabbit.
ORGANISM      Oryctolagus cuniculus
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
REFERENCE     1 (bases 1 to 234)
AUTHORS       Sakai, D., Tong, H.-S. and Minklin, C.
TITLE         Osteoclast Molecular Phenotyping by Random cDNA Sequencing
JOURNAL       Bone 17 (2), 111-119 (1995)
MEDLINE       96021365
COMMENT       Other ESTs: RABEST038T, RABEST045T, RABEST084T, RABEST103T,
               RABEST140T, RABEST202T, RABEST229T
               Contact: Sakai D
               Basic Sciences
               University of Southern California
               USC School of Dentistry, 925 West 34th Street, DEN-4220, Los
               Angeles, CA 90089-0641
               Tel: 2137405563
               Fax: 2137407560

FEATURES
source
Email: sakai@molbio.usc.edu
Seq primer: T7 promoter.
Location/Qualifiers
1..234
/organism="Oryctolagus cuniculus"
/strain="New Zealand White"
/db_xref="taxon:9986"
/clone="PRABOC163"
/clone_lib="Rabbit Osteoclast, Dennis Sakai"
/lab_host="E. coli DH12S"
/note="Vector: pSPORT1; Site_1: SalI; Site_2: NotI; Poly(A)
)+ RNA was purified from a 97% pure population of
osteoclasts prepared from the long bones of 10 day old
rabbits. First strand cDNA was synthesized by priming
with an oligo(dT)-NotI anchor-primer and second strand
cDNA was synthesized by replacement synthesis as described
by Gubler and Hoffman (Gene 25:283, 1983). Following the
addition of SalI adapters and NotI digestion, the cDNA was
cloned between the SalI (50) and NotI (30) sites of the
pSPORT1 (BRL) plasmid vector."
BASE COUNT    57 a 62 c 81 g 34 t
ORIGIN
Query Match    20.1%; Score 43; DB 14; Length 234;
Best Local Similarity 54.8%; Pred. No. 6.8;
Matches 85; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

QY 5 GGAGAGCGCTACGACCTGGCCCGCGTGGCGCGGTACAAAGGTGAACAAGAGCTGGTCT 64
    || || || || || || || || || || || || || || || || || || ||
Db 12 GAAGAAGCGCAAGAGGCGGTCTGTCTGCCTGAGGAGGACGAAGACATCGTGT 71
    || || || || || || || || || || || || || || || || || || ||
QY 65 TGGCGGTGCCCAACCGGCTCTGTGTGACTGCCACCGCTCACCGAGGAAGACGTGTCGC 124
    || || || || || || || || || || || || || || || || || || ||
Db 72 GGAGGAGGGCAAGAGAGATCTGTGTGGCGGACGTGGGCCAGACCGTGGAGGACCCCTACGC 131
    || || || || || || || || || || || || || || || || || || ||
QY 125 CACCATCGGGTACTGTGTGCGCGCTGCACGAGGGCC 159
    || || || || || || || || || || || || || || || || || || ||
Db 132 CACCTTCGTGCATGCTGCCCGCACGAAGGACTGCC 166
    || || || || || || || || || || || || || || || || || || ||

RESULT 10
R86566
LOCUS          R86566                235 bp      mRNA      linear      EST 17-AUG-1995
DEFINITION    RABEST140T Rabbit Osteoclast, Dennis Sakai Oryctolagus cuniculus
               cDNA clone PRABOC140 5' similar to cofillin, mRNA sequence.
ACCESSION     R86566
VERSION       R86566.1 GI:947220
KEYWORDS      EST.
SOURCE        rabbit.
ORGANISM      Oryctolagus cuniculus
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
REFERENCE     1 (bases 1 to 235)
AUTHORS       Sakai, D., Tong, H.-S. and Minklin, C.
TITLE         Osteoclast Molecular Phenotyping by Random cDNA Sequencing
JOURNAL       Bone 17 (2), 111-119 (1995)
MEDLINE       96021365
COMMENT       Other ESTs: RABEST038T, RABEST045T, RABEST084T, RABEST103T,
               RABEST163T, RABEST202T, RABEST229T
               Contact: Sakai D
               Basic Sciences
               University of Southern California
               USC School of Dentistry, 925 West 34th Street, DEN-4220, Los
               Angeles, CA 90089-0641
               Tel: 2137405563
               Fax: 2137407560
               Email: sakai@molbio.usc.edu
               Seq primer: T7 promoter.
               Location/Qualifiers
               1..235
               /organism="Oryctolagus cuniculus"
               /strain="New Zealand White"
               /db_xref="taxon:9986"

```

```

Query Match      20.1%   Score 43;   DB 14;   Length 331;
Best Local Similarity 54.8%   Pred. No. 7.1;
Matches 85;   Conservative 0;   Mismatches 70;   Indels 0;   Gaps 0;

```

	Query Match	20.1%	Score 43	DB 14	Length 331
	Best Local Similarity	54.8%	Pred. No. 7.1		
	Matches 85	Conservative 0	Mismatches 70	Indels 0	Gaps 0
QY	5	GGAGAACGCTACGACCTGGCCCGCTGGCCCGGTACAAGTGAACAAAGACGTGGGTCT	64		
Db	12	GAAGACGCAAGAGCGGTGTTCTCCCTGAGGAGGCAAGAAACATCATGTGT	71		
QY	65	TGGCGGTGCCAACCCGCGCTCTGGTACTGCCACACGCTCACCGAGGAAGACGTCGTGCG	124		
Db	72	GGAGGAGGCAAGAGATCTGTTGGCGACGTTGGGCCACCGTGGAGGACCCCTACGC	131		
QY	125	CACCATCGGTACTTGGTGGCCCTGCACGAGGCC	159		
Db	132	CACCTTGTGCATGCTGCCCAAGGACTGCC	166		

RESULT	12
LOCUS	BF264852
DEFINITION	BF264852 878 bp mRNA linear EST 23-OCT-2001 HV_Cea0010124f Hordeum vulgare seedling green leaf EST library HVCNA0004 (Blumeria challenged) Hordeum vulgare cDNA clone HV_Cea0010124f, mRNA sequence.

LOCUS	BF264852	878 bp	mRNA	linear	EST 23-OCT-2001
DEFINITION	HV_CEA0010124f Hordeum vulgare seedling green leaf EST library				
	HV_CEA0010004 (Blumeria challenge) Hordeum vulgare cDNA clone				
	HV_CEA0010124f, mRNA sequence.				
ACCESSION	BF264852				
VERSION	BF264852.2 GI:13261763				
KEYWORDS	EST.				

REFERENCE	TITLE	JOURNAL	COMMENT
1 (bases 1 to 878)			
Wing, R., Close, T.J., Kleinhofs, A., Wise, R., Wei, F., Begum, D., Fritsch, D., Yu, Y., Henry, D., Palmer, M., Rambo, T., Simmons, J., Choi, D.W., Fenton, R.D., Oates, R. and Main, D.	Development of a genetically and physically anchored EST resource for barley genomics: Blumeria infected incompatible (Mla13) seedling leaf cDNA library	Unpublished (2001)	On Nov 17, 2000 this sequence version replaced gi:11195846.
Contact: Wing RA			
Clemson University Genomics Institute			
Clemson University			
100 Jordan Hall, Clemson, SC 29634, USA			
Tel: 864 656 7288			
Fax: 864 656 4293			
Email: rwing@clemson.edu			
Total hg bases = 224			
Seq primer: AATTAACCTCTACTAAGGG			
High quality sequence stop: 348.			



) of *Blumeria graminis* f. sp. *hordei*, and leaves were harvested 20 and 24 hr post-inoculation and snap frozen; uninoculated leaves were harvested 20 hr post-inoculation (Wei, Wise). In the TJ Close lab at the University of California, Riverside, total RNA was prepared from each sample pool, equal quantities of all three RNA pools were combined, poly(A) RNA was purified from the mixture, one cDNA library was made, and 1 million pfu were in vivo excised to give pBluescript SK(-) cDNA phagemids (Choi, Close). Phagemids were plated and picked at the Clemson University Genomics Institute (CUGI) (Begum, Palmer, Frisch, Atkins and Wing). Plasmid DNA preparations, DNA sequencing and sequence analysis were performed at CUGI (Wing, Yu, Frisch, Henry, Simmons, Oates, Rambo, Main). The sequence has been trimmed to remove vector sequence and contains a minimum of 100 bases of phred value 20 or above. For more details on library preparation and sequence analysis see <http://www.genome.clemson.edu/projects/barley>. To order this clone see <http://www.genome.clemson.edu/orders> Also see Close TJ, Wing R, Kleinhofs A, Wise R (2001) Genetically and physically anchored EST resources for barley genomics. *Barley Genetics Newsletter* 31:29-30. (<http://wheat.pw.usda.gov/ggpages/bgn/31/cover.html>)

1. .618  
/organism="Triticum aestivum"  
/cultivar="Chinese Spring"  
/db\_xref="taxon:4565"  
/clone="who5m04"  
/clone\_lib="Y. Ogihara unpublished cDNA library, Wh\_oh"  
/tissue\_type="pistil at heading date"  
/dev\_stage="Feekes' scale 10.5"  
/note="Vector: Lambda Uni-ZAP XR, excised phagemid; Site 1: EcoRI; Site 2: XhoI; Plants were grown under hydroponic conditions at UC Davis, salt stressed for 12 hours, and for 7 days, then dissected and frozen (Akhunov in J Dvorak Lab). Total RNA was prepared from sheath tissue, equal quantities of RNA were pooled from the two samples, polyA was purified from the pooled RNA, a cDNA library was made, and the cDNA clones were in vivo excised to give pBluescript phagemids in the TJ Close lab at the University of California, Riverside (Akhunov, Chin, Choi, Close, Fenton, Kiantan, Otto, Simons, Zhang). Plasmid DNA preparations and DNA sequencing were performed in the OD Anderson lab (all other authors)."

BASE COUNT 205 a 152 c 288 g 232 t 1 others  
ORIGIN  
Query Match 20.1%; Score 43; DB 12; Length 878;  
Best Local Similarity 50.7%; Pred. No. 8;  
Matches 103; Conservative 0; Mismatches 100; Indels 0; Gaps 0;  
QY 5 GGAGAGCGCTACGACCTGGCGCGGTACAGGTGACAAAGCTGGTCT 64  
Db 93 GGACTAGCGAGGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 152  
QY 65 TGGCGGTGCCAACCGCGCTGTGTACTGCCACACGCTCACCGAGGAGAGCTGCTGC 124  
Db 153 CGGGTGGAGAACCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 212  
QY 125 CACCATCGGTACTGTGGTGGCTGCACGAGGCCAGACGATGACCGCCCGCGCGG 184  
Db 213 CGCGCAGCGGTTCAGTGGATCAAGGGCAAGGTCCAAAAGAGACGAGGCGCGCGG 272  
QY 185 CTTCTGAGGTCCCGGTGAGGTGCG 207  
Db 273 CGCGCGGCGACTAGCGCGCTGTGCG 295

RESULT 13  
BJ270031  
LOCUS BJ270031 618 bp mRNA linear EST 09-APR-2002  
DEFINITION BJ270031 Y. Ogihara unpublished cDNA library, Wh\_oh Triticum aestivum cDNA clone who5m04 5', mRNA sequence.  
ACCESSION BJ270031  
VERSION BJ270031  
KEYWORDS EST.  
SOURCE bread wheat.  
ORGANISM Triticum aestivum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae  
; Triticeae; Triticum.  
REFERENCE 1 (bases 1 to 618)  
AUTHORS Ogihara, Y. and Murai, K.  
TITLE Expressed genes in Triticum aestivum  
JOURNAL Unpublished (2002)  
COMMENT Contact: Tadasu Shin-i  
Center For Genetic Resource Information  
National Institute of Genetics  
1111 Yata, Mishima, Shizuoka 411-8540, Japan  
Tel: 81-559-81-6856  
Fax: 81-559-81-6855  
Email: [tshini@genes.nig.ac.jp](mailto:tshini@genes.nig.ac.jp).  
Location/Qualifiers

BASE COUNT 113 a 214 c 207 g 84 t  
ORIGIN  
Query Match 19.9%; Score 42.6; DB 13; Length 618;  
Best Local Similarity 50.2%; Pred. No. 9.4;  
Matches 105; Conservative 0; Mismatches 104; Indels 0; Gaps 0;  
QY 2 CAAGGAGAGCGCTACGACCTGGCGCGGTGGCGGTGACAAAGAGAGCTGGG 61  
Db 209 CAACGCCGAGGTCTACGCCACGCGCGCGCGGAGAGATCATGGGCGGCGCATCCGGGA 268  
QY 62 TCTTGGCGGTGCCAACCGCGCTCTGTGTACTGCCACACGCTCACCGAGGAGAGCTCGT 121  
Db 269 CTGTGGCTGGCGCGCGCGCGCGCGCTCATCTCCACCAAGCTCTCTGGGCGCGCAGGG 328  
QY 122 CGCCACCATCGGTACCTGTGGCGCTGCACGAGGCCAGACGATGACCGCCCGCGG 181  
Db 329 CCCCAACGACGAGGCGCTCTCCCGCAAGCACATCGTGAGGGCGCTCAAGGCGCTCGCTCAA 388  
QY 182 CGGCTCGAGGTCCCGGTGAGGTGCGAGTGCAGC 210  
Db 389 CGGCTCAACATGGAGTACGTGACGTGC 417  
RESULT 14  
BJ226380  
LOCUS BJ226380 623 bp mRNA linear EST 05-APR-2002  
DEFINITION BJ226380 Y. Ogihara unpublished cDNA library, Wh\_dL Triticum aestivum cDNA clone whdL26d22 5', mRNA sequence.  
ACCESSION BJ226380  
VERSION BJ226380  
KEYWORDS EST.  
SOURCE bread wheat.  
ORGANISM Triticum aestivum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae  
; Triticeae; Triticum.  
REFERENCE 1 (bases 1 to 623)  
AUTHORS Ogihara, Y. and Murai, K.  
TITLE Expressed genes in Triticum aestivum  
JOURNAL Unpublished (2002)  
COMMENT Contact: Tadasu Shin-i  
Center For Genetic Resource Information  
National Institute of Genetics  
1111 Yata, Mishima, Shizuoka 411-8540, Japan  
Tel: 81-559-81-6856  
Fax: 81-559-81-6855  
Email: [tshini@genes.nig.ac.jp](mailto:tshini@genes.nig.ac.jp).  
Location/Qualifiers  
1. .623  
/organism="Triticum aestivum"  
/cultivar="Chinese Spring"

```

/db_xref="taxon:4565"
/clone="whd126d22"
/clone_lib="Y. Ogihara unpublished cDNA library, wh_dL"
/tissue_type="crown of seedling"
/dev_stage="freekes' scale 1"
/notes="Vector: Lambda Uni-ZAP XR, excised phagemid;
Site.1: EcoRI; Site.2: XhoI; Plants were grown under
hydroponic conditions at UC Davis, salt stressed for 12
hours, and for 7 days, then dissected and frozen (Akhunov
in J. Dvorak Lab). Total RNA was prepared from sheath
tissue, equal quantities of RNA were pooled from the two
samples, polyA was purified from the pooled RNA, a cDNA
library was made, and the cDNA clones were in vivo
excised to give pBluescript phagemids in the TJ Clouse lab
at the University of California, Riverside (Akhunov, Chin
, Choi, Clouse, Fenton, Kianian, Otto, Simons, Zhang).
Plasmid DNA preparations and DNA sequencing were
performed in the OD Anderson lab (all other authors)."
BASE COUNT      114 a      208 g      84 t
ORIGIN
Query Match      19.9%; Score 42.6; DB 13; Length 623;
Best Local Similarity 50.2%; Pred. No. 9.4;
Matches 105; Conservative 0; Mismatches 104; Indels 0; Gaps 0;

QY 2 CAAGGAGAACGCTACGACCTGGCCCGCTGCGGTACAAGGTCAACAAGAGCTGGG 61
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 214 CAACGCCGAGGTCTACGCCAACGCCGCGCGGAGGAGATCATGGGCGCGCATCCGGGA 273

QY 62 TCTTGGGGTGCCAACCCGGCTCTGTGTACTGCCACCGCTACCCGAGGAGAGCTGGT 121
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 274 CTTGGGCTGGCGCGCGGAGGTCTGTATCTCCACCAAGCTCTTCTGGGCGGCCAGGG 333

QY 122 CGCACCATCGGTACTGTGGCTGCTGCACGAGGCGGACACGATGACCGCCCCCGG 181
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 334 CCCACGACAGGGGCTCTCCCGAAGCATCTCGAGGGGCTCAAGGGCTCGCTCAA 393

QY 182 CGGCTCGAGGTCCCGTCTGAGGTGACG 210
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 394 GCGGCTCAACATGGAGTACGTCGACGTCG 422

```

```

RESULT 15
BE361556
LOCUS      DGI_72_E07.b1_A002 Dark Grown 1 (DGI) Sorghum bicolor cDNA, mRNA
DEFINITION sequence.
ACCESSION BE361556
VERSION   BE361556.1 GI:9303015
KEYWORDS  EST.
SOURCE    sorghum.
ORGANISM  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade: Panicoideae; Andropogoneae; Sorghum.
1 (bases 1 to 479)
Cordonnier-Pratt,M.-M., Gingle,A., Marsala,C., Sudman,M. and Pratt
,L.H.
An EST database from Sorghum: dark-grown seedlings
Unpublished (2000)
Contact: Cordonnier-Pratt MM
Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 583 0210
Email: mmpratt@uga.edu
Sequences have been trimmed to exclude PolyA, vector and regions
below Phred quality 16. The threshold for highest quality sequence
is 20.
Seq primer: JEN REV
High quality sequence stop: 456
POLYA-No.

```

```

FEATURES
    source              Location/Qualifiers
     ..479
     /organism="Sorghum bicolor"
     /db_xref="taxon:4558"
     /clone_lib="Dark Grown 1 (DGI)"
     /note="Organ: 5-day-old dark-grown seedlings; Vector:
     Lambda zap; Site.1: XhoI; Site.2: EcoRI; The library was
     made from poly-A RNA in the cloning vector Lambda ZAP II.
     Clones to be sequenced were prepared by mass excision."
BASE COUNT      85 a      158 c      170 g      66 t
ORIGIN
Query Match      19.8%; Score 42.4; DB 10; Length 479;
Best Local Similarity 50.5%; Pred. No. 10;
Matches 103; Conservative 0; Mismatches 101; Indels 0; Gaps 0;

QY 11 GCGCTACGACCTGGCCCGCTGGCGGCTACAAGGTGAACAAGAAGCTGGTCTTTGGCGG 70
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 214 GCCCAAGGCCGTGTGTCAGGTGCTGCAAGACACACGGCATCAAGAAGGTGAAGCTTTTCGA 273

QY 71 TGCCAACCCGGCTCTGTGTACTGCGCACACGCTCACCGAGGAAGACGCTCGTCGCCACCAT 130
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 274 CACCGACCCGCGCCCATGACGCGCTCGCGCGGACCCGGCATCGAGGTCTATGTCGCCCAT 333

QY 131 CGGGTACTGTGTGGCTGTCACGAGGGCCAGACCATGACCCCCCGGCGGCTCGA 190
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 334 CCCCAACAACATGTCTCGCGGACCTCGCGGAGCGAGCGGAAAGGCTAAGGACTGGGTCAA 393

QY 191 GGTCCCGGTGAGGTGCGACGACAT 214
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 394 GCGCAACGTCCGGCGCTACGACTT 417

```

Search completed: November 13, 2002, 03:59:41  
Job time : 1156.63 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 12, 2002, 16:05:40 : Search time 21.7711 Seconds  
(without alignments)  
3487.380 Million cell updates/sec

Title: US-09-697-123b-7

Perfect score: 214

Sequence: 1 tcaaggagagcgtacgac.....ccggtcgaggtcgacgacat 214

Scoring table: IDENTITY\_NUC

Gapop 10.0, Gapext 1.0

Searched: 320260 seqs, 177392727 residues

Total number of hits satisfying chosen parameters: 640520

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications\_NA:\*

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*
- 12: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*
- 13: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*
- 14: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Query %	Score	Match	Length	DB ID	Description
1	88.4	41.3	5096	10	US-09-984-711-5	Sequence 5, Appl1
2	88.4	41.3	5099	9	US-10-075-460-5	Sequence 5, Appl1
3	88.4	41.3	5099	10	US-09-887-052-1	Sequence 1, Appl1
4	88.4	41.3	5099	10	US-09-887-052-3	Sequence 3, Appl1
5	88.4	41.3	5099	10	US-09-887-052-5	Sequence 5, Appl1
6	40	18.7	3343	10	US-09-950-772-7	Sequence 7, Appl1
7	40	18.7	15872	10	US-09-861-289-1	Sequence 1, Appl1
8	38	17.8	1380	10	US-09-950-772-3	Sequence 3, Appl1
9	37.6	17.6	2898	10	US-09-737-149-3	Sequence 3, Appl1
10	36.6	17.1	584	10	US-09-867-550-669	Sequence 669, App
11	36.2	16.9	221	10	US-09-864-761-26660	Sequence 26660, A
12	36.2	16.9	392	10	US-09-878-574-3598	Sequence 3598, App
13	36.2	16.9	1914	10	US-09-815-242-7960	Sequence 7960, Ap
14	36.2	16.9	3057	10	US-09-815-242-4131	Sequence 4131, Ap
15	35.6	16.6	242	10	US-09-923-876-2475	Sequence 2475, Ap
16	35.6	16.6	244	10	US-09-923-876-2527	Sequence 2527, Ap
17	35.6	16.6	248	10	US-09-923-876-1166	Sequence 1166, Ap
18	35.6	16.6	248	10	US-09-923-876-2335	Sequence 2335, Ap
19	35.6	16.6	260	10	US-09-923-876-2086	Sequence 2086, Ap

#### ALIGNMENTS

##### RESULT 1

US-09-984-711-5  
; Sequence 5, Application US/09984711  
; Patent No. US20020119549A1  
; GENERAL INFORMATION:  
; APPLICANT: MOECKEL, Bettina  
; APPLICANT: BATHE, Brigitte  
; APPLICANT: STEPHAN, Hans  
; APPLICANT: KREUTZER, Caroline  
; APPLICANT: HERMANN, Thomas  
; APPLICANT: PFEFFERLE, Walter  
; APPLICANT: BINDER, Michael  
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES WHICH CODE FOR THE rpsL GENE  
; FILE REFERENCE: 204209US0  
; CURRENT APPLICATION NUMBER: US/09/984,711  
; CURRENT FILING DATE: 2001-10-31  
; PRIOR APPLICATION NUMBER: DE10108230.9  
; PRIOR FILING DATE: 2001-02-16  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 5  
; LENGTH: 5096  
; TYPE: DNA  
; ORGANISM: Corynebacterium glutamicum  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (702)..(4196)  
; OTHER INFORMATION:  
US-09-984-711-5

Query Match 41.3%; Score 88.4; DB 10; Length 5096;

Best Local Similarity 68.4%; Pred. No. 2.7e+13;

Matches 141; Conservative 0; Mismatches 56; Indels 9; Gaps 1;

QY 9 AAGCGGTACGACCTGCCCGCGTGGCGGTACAAAGTGAACAAAGCTGGGTCTTGGC 68

Db 1578 AAGCGGTACGACCTGCCCGTGGTGTGCGTTACAGATCAACCGCAAGCTCGSCCTTGGT 1637

QY 69 GGTGCCAACCCGCTCTGGTGACTGCCACCTCACCAGGAGAACGCTGCTGCCACC 128

Db 1638 GCGGACCAACGATGTTTGTATGACT-----CTTACTGAGAGGACATCGCAACACC 1688

20 35.6 16.6 266 10 US-09-923-876-2207 Sequence 2207, Ap  
21 35.6 16.6 312 10 US-09-923-876-3681 Sequence 3681, Ap  
22 35.6 16.6 2916 10 US-09-737-149-5 Sequence 5, Appl1  
23 35.6 16.6 3132 10 US-09-923-876-2224 Sequence 7, Appl1  
24 35.2 16.4 235 10 US-09-923-876-2583 Sequence 2583, Ap  
25 35.2 16.4 254 10 US-09-815-242-7982 Sequence 7982, Ap  
26 35.2 16.4 1434 10 US-09-815-242-7989 Sequence 7989, Ap  
27 35.2 16.4 1818 10 US-09-815-242-7873 Sequence 7873, Ap  
28 34.8 16.3 888 10 US-09-880-107-3702 Sequence 3702, Ap  
29 34.6 16.2 446 10 US-09-732-680A-1 Sequence 1, Appl1  
30 34.4 16.1 1539 10 US-09-957-974-2 Sequence 3537, Ap  
31 34.4 16.1 12606 10 US-09-923-876-3537 Sequence 320, App  
32 34.2 16.0 239 10 US-09-954-456-320 Sequence 320, App  
33 34.2 16.0 1713 10 US-09-897-214-13 Sequence 13, Appl1  
34 34.2 16.0 2223 10 US-09-864-761-28400 Sequence 28400, A  
35 34 15.9 215 10 US-09-923-876-3118 Sequence 3118, Ap  
36 34 15.9 224 10 US-09-923-876-3118 Sequence 3118, Ap  
37 34 15.9 229 10 US-09-864-761-11830 Sequence 11830, A  
38 34 15.8 660 10 US-09-815-242-7875 Sequence 7875, Ap  
39 34 15.8 1476 10 US-09-997-664-99 Sequence 99, Appl1  
40 33.8 15.8 1792 10 US-09-815-242-7772 Sequence 7772, Ap  
41 33.8 15.8 6491 10 US-09-997-664-1 Sequence 1, Appl1  
42 33.6 15.7 258 10 US-09-923-876-1695 Sequence 1695, Ap  
43 33.6 15.7 939 10 US-09-822-830A-494 Sequence 494, App  
44 33.6 15.7 939 10 US-09-822-830A-494 Sequence 494, App  
45 33.6 15.7 939 10 US-09-822-830A-494 Sequence 494, App

```
QY 129 ATCGGTACCTGCTGGTCCGCTGCAGGAGGCGCCAGACCATGACCGCCCGCGGCGCTC 188
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1689 ATCGACTACTGCTGGCTGTCACCGAGTGAGCCGCTGATGACTTCTCCAATGGTGAA 1748
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 189 GAGGTCCCGGTGAGGTGAGACAT 214
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1749 GAGATCCCATGTCGAGACCGATGACAT 1774
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 2
US-10-075-460-5
: Sequence 5, Application US/10075460
: Patent No. US2002015557A1
: GENERAL INFORMATION:
: APPLICANT: MOCKEL, BETTINA
: APPLICANT: BATHE, BRIGITTE
: APPLICANT: HANS, STEFAN
: APPLICANT: KREUTZER, CAROLINE
: APPLICANT: HERMANN, THOMAS
: APPLICANT: PFEFFERLE, WALTER
: APPLICANT: BINDER, MICHAEL
: TITLE OF INVENTION: NUCLEOTIDE SEQUENCES WHICH CODE FOR THE rpsL GENE
: FILE REFERENCE: 218472USOX
: CURRENT APPLICATION NUMBER: US/10/075.460
: PRIOR FILING DATE: 2002-02-15
: PRIOR APPLICATION NUMBER: DE 10107230.9
: PRIOR FILING DATE: 2001-02-16
: PRIOR APPLICATION NUMBER: DE 10162386.0
: PRIOR FILING DATE: 2001-12-19
: NUMBER OF SEQ ID NOS: 14
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 5
: LENGTH: 5099
: TYPE: DNA
: ORGANISM: Corynebacterium glutamicum
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (702)..(4196)
: OTHER INFORMATION:
US-10-075-460-5

Query Match 41.3%; Score 88.4; DB 9; Length 5099;
Best Local Similarity 68.4%; Pred. No. 2.7e-13;
Matches 141; Conservative 0; Mismatches 56; Indels 9; Gaps 1;

QY 9 AAGCCTACGACCTGGCCGCGTGGCGGTACAAAGGTGAACAAGAGCTGGGTCTTGGC 68
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1578 AAGCGCTACGACCTGGCTCGCGTTGGTGTACAAAGATCAACCGCAAGCTCGGCCTTGGT 1637
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 69 GGTGCCAACCCGGCTCTGGTGACTGCCACACGCTCACCGAGGAGACGTGCTGCCACC 128
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1638 GCGCACCATGATGTTTGATGACT-----CTTACTGAAGAGGACATCGCAACCACC 1688
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 129 ATCGGTACCTGCTGGCTGCGCTGCAGAGGCGCCAGACCATGACCGCCCGCGGCGCTC 188
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1689 ATCGAGTACCTGCTGGCTGCTGCACGAGGTGAGCGCGTCTGCTCTCCAAATGGTGAA 1748
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 189 GAGGTCCCGGTGAGGTGCGACGACAT 214
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1749 GAGATCCCATGTCGAGACCGATGACAT 1774
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 4
US-09-887-052-3
: Sequence 3, Application US/09887052
: Patent No. US20020119537A1
: GENERAL INFORMATION:
: APPLICANT: MOECKEL, Bettina
: APPLICANT: BATHE, Brigitte
: APPLICANT: HERMANN, Thomas
: APPLICANT: PFEFFERLE, Walter
: APPLICANT: BINDER, Michael
: TITLE OF INVENTION: NUCLEOTIDE SEQUENCES WHICH CODE FOR THE rpoB GENE
: FILE REFERENCE: 204212USOX
: CURRENT APPLICATION NUMBER: US/09/887,052
: PRIOR FILING DATE: 2001-06-25
: PRIOR APPLICATION NUMBER: DE10107229.5
: PRIOR FILING DATE: 2001-02-16
: NUMBER OF SEQ ID NOS: 8
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 3
: LENGTH: 5099
: TYPE: DNA
: ORGANISM: Corynebacterium glutamicum
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (702)..(4196)
: OTHER INFORMATION:
US-09-887-052-3

Query Match 41.3%; Score 88.4; DB 10; Length 5099;
Best Local Similarity 68.4%; Pred. No. 2.7e-13;
Matches 141; Conservative 0; Mismatches 56; Indels 9; Gaps 1;

QY 9 AAGCCTACGACCTGGCCGCGTGGCGGTGAACAAGGTGAACAAGAGCTGGGTCTTGGC 68
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1578 AAGCGCTACGACCTGGCTCGCGTTGGTGTACAAAGATCAACCGCAAGCTCGGCCTTGGT 1637
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 69 GGTGCCAACCCGGCTCTGGTGACTGCCACACGCTCACCGAGGAGACGTGCTGCCACC 128
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1638 GCGCACCATGATGTTTGATGACT-----CTTACTGAAGAGGACATCGCAACCACC 1688
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 129 ATCGGTACCTGCTGGCTGCGCTGCAGAGGCGCCAGACCATGACCGCCCGCGGCGCTC 188
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1689 ATCGAGTACCTGCTGGCTGCTGCACGAGGTGAGCGCGTCTGCTCTCCAAATGGTGAA 1748
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 189 GAGGTCCCGGTGAGGTGCGACGACAT 214
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1749 GAGATCCCATGTCGAGACCGATGACAT 1774
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 3
US-09-887-052-1
: Sequence 1, Application US/09887052
: Patent No. US20020119537A1
: GENERAL INFORMATION:
: APPLICANT: MOECKEL, Bettina
: APPLICANT: BATHE, Brigitte
: APPLICANT: HERMANN, Thomas
: APPLICANT: PFEFFERLE, Walter
: APPLICANT: BINDER, Michael
: TITLE OF INVENTION: NUCLEOTIDE SEQUENCES WHICH CODE FOR THE rpoB GENE
```

Db 1638 GCCGACCACGATGTTTGATGACT-----CTTACTGAAGAGGACATCGCAACCAACC 1688  
Qy 129 ATCGGGTACCTGTGGCCCTGCGACGAGGCGCCAGACCATGACCGCCCGCGGCGCTC 188  
Db 1689 ATCAGGTACCTGGTGGCTCTGCACGCGAGGTGAGCGGCTCATGACTTCTCCAAATGCTGAA 1748  
Qy 189 GAGGTCCCGGTGCGAGGTGCGAGCAT 214  
Db 1749 GAGATCCCACTGCGAGACCGATGACAT 1774

RESULT 5  
US-09-887-052-5  
; Sequence 5, Application US/09887052  
; Patent No. US20020119537A1  
; GENERAL INFORMATION:  
; APPLICANT: MOECKEL, Bettina  
; APPLICANT: BATHE, Brigitte  
; APPLICANT: HERMANN, Thomas  
; APPLICANT: PFEFFERLE, Walter  
; APPLICANT: BINDER, Michael  
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES WHICH CODE FOR THE IPoB GENE  
; FILE REFERENCE: 204212US0X  
; CURRENT APPLICATION NUMBER: US/09/887,052  
; CURRENT FILING DATE: 2001-06-25  
; PRIOR APPLICATION NUMBER: DE10107229.5  
; PRIOR FILING DATE: 2001-02-16  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 5  
; LENGTH: 5099  
; TYPE: DNA  
; ORGANISM: Corynebacterium glutamicum  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (702)..(4196)  
US-09-887-052-5

Query Match 41.3%; Score 88.4; DB 10; Length 5099;  
Best Local Similarity 68.4%; Pred. No. 2.7e-13;  
Matches 141; Conservative 0; Mismatches 56; Indels 9; Gaps 1;  
Qy 9 AAGCGCTACGACCTGGCGCGGTGACAAAGGTGAACAAAGTGGGTCTTTGGC 68  
Db 1578 AAGCGCTACGACCTGGCGCGGTGCTGTACAAAGTCAACCGCAAGCTCGCCCTTGGT 1637  
Qy 69 GGTGCCAACCGCGCTCTGGTACTGCCACACCGCTCACCGAGGAAGACGTGCTGCCACCC 128  
Db 1638 GCGCACCACGATGTTTGATGACT-----CTTACTGAAGAGGACATCGCAACCAACC 1688  
Qy 129 ATCGGGTACCTGTGGCGCTGCGACGAGGCGCCAGACCATGACCGCCCGCGGCGCTC 188  
Db 1689 ATCAGGTACCTGGTGGCTCTGCGACGAGGTGAGCGGCTCATGACTTCTCCAAATGCTGAA 1748  
Qy 189 GAGGTCCCGGTGCGAGGTGCGAGCAT 214  
Db 1749 GAGATCCCACTGCGAGACCGATGACAT 1774

RESULT 6  
US-09-950-772-7  
; Sequence 7, Application us/09950772  
; Patent No. US20020102713A1  
; GENERAL INFORMATION:  
; APPLICANT: SUZUKI, Shunichi  
; APPLICANT: ONISHI, No. US2002010713Allimasa  
; APPLICANT: YOKOZEKI, Kenzo  
; TITLE OF INVENTION: 5-SUBSTITUTED HYDANTOIN RACEMASE, DNA CODING FOR THE RACEMASE, AND  
; FILE REFERENCE: 212318US05  
; CURRENT APPLICATION NUMBER: US/09/950,772  
; CURRENT FILING DATE: 2001-09-13

; PRIOR APPLICATION NUMBER: JP2000-278571  
; PRIOR FILING DATE: 2000-09-13  
; PRIOR APPLICATION NUMBER: JP2001-65815  
; PRIOR FILING DATE: 2001-03-08  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 7  
; LENGTH: 3343  
; TYPE: DNA  
; ORGANISM: Microbacterium liquefaciens  
US-09-950-772-7

Query Match 18.7%; Score 40; DB 10; Length 3343;  
Best Local Similarity 51.1%; Pred. No. 0.13;  
Matches 94; Conservative 0; Mismatches 90; Indels 0; Gaps 0;  
Qy 28 GCGTGGCGCGGTACAAAGGTGAACAAAGACTGGGTCTTGGCGGTGCCAACCGGCTCTGG 87  
Db 721 GCGATCTCATGTTTCGACGCTCATTTGGAAGAACTGTCGAGTGGTTCCAGTCAAGGCGCATCA 780  
Qy 88 TGACTGCCACACGCTCACCGAGGAGGAGAGTGTGTCGCCACCATCGGGTACCTGTGTGCC 147  
Db 781 TCGAAGCCGACATCTCTGTGAAGGACGCGCGGATCGCCGCCATCAGCGAGGAGCCCTCG 840  
Qy 148 TGCACGAGGCGCCAGACCATGATGACCGCCCGCGGCTCGAGGTCCCGGTGCGAGGTGCG 207  
Db 841 AGCGCGAAGCCCGCGGACCATCATGATGCGCAGGAGGTTCGTGATGCCCGGTGGTTCG 900  
Qy 208 ACGA 211  
Db 901 ATGA 904

RESULT 7  
US-09-861-289-1  
; Sequence 1, Application US/09861289  
; Patent No. US20020110897A1  
; GENERAL INFORMATION:  
; APPLICANT: Sherman, D.H.  
; APPLICANT: Liu, H.  
; APPLICANT: Xue, Y.  
; APPLICANT: Zhao, L.  
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin  
; FILE REFERENCE: 600.438USI  
; CURRENT APPLICATION NUMBER: US/09/861,289  
; CURRENT FILING DATE: 2001-05-18  
; PRIOR APPLICATION NUMBER: 09/105,537  
; PRIOR FILING DATE: 1998-06-26  
; NUMBER OF SEQ ID NOS: 43  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1  
; LENGTH: 15872  
; TYPE: DNA  
; ORGANISM: Streptomyces venezuelae  
US-09-861-289-1

Query Match 18.7%; Score 40; DB 10; Length 15872;  
Best Local Similarity 53.1%; Pred. No. 0.15;  
Matches 85; Conservative 0; Mismatches 75; Indels 0; Gaps 0;  
Qy 53 GAAGCTGGGTCTTGGCGGTGCCAACCCGGCTCTGTGTGACTGCCACCATCGCTCACCGAGGA 112  
Db 2416 GGAGCTCGGTCCCGACGGGTCTGTCCCGATGCGCGGACTCCGTACGCGACAGGA 2475  
Qy 113 AGACGTGCTGCGCCACCATCGGGTACTGTGCGCTGACAGGAGGCGCCAGACCATCATGAC 172  
Db 2476 GCGCGCCAGCGGTCTCGCCCTCGGCAAGGCGCCCGGAGGCCAGTCTGCTGCTGCTCGC 2535  
Qy 173 CGCCCGCGCGGCGCTCGAGGTCCCGGTTCGAGTTCGACGAC 212  
Db 2536 CGCACTCCACCGCTTCTCGTCCGGGCGCACACGCTCGAC 2575

```
RESULT 8
US-09-950-772-3
; Sequence 3, Application US/09950772
; Patent No. US20020102713A1
; GENERAL INFORMATION:
; APPLICANT: SUZUKI, Shunichi
; APPLICANT: ONISHI, No. US20020102713Alimasa
; APPLICANT: YOKOZEKI, Kenzo
; TITLE OF INVENTION: 5-SUBSTITUTED HYDANTOIN RACEMASE, DNA CODING FOR THE RACEMASE, AND
; FILE REFERENCE: 212318USO
; CURRENT APPLICATION NUMBER: US/09/950,772
; CURRENT FILING DATE: 2001-09-13
; PRIOR APPLICATION NUMBER: JP2000-278571
; PRIOR FILING DATE: 2000-09-13
; PRIOR APPLICATION NUMBER: JP2001-65815
; PRIOR FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1380
; TYPE: DNA
; ORGANISM: Microbacterium liquefaciens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1377)
; OTHER INFORMATION:
US-09-950-772-3

Query Match      17.8%; Score 38; DB 10; Length 1380;
Best Local Similarity 51.1%; Pred. No. 0.38;
Matches 89; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

QY 38 GTACAAGGTGAACAAGAGCTGGTCTTTGGCGGTGCCAACCCGGCTCTGGTACTGCCAC 97
DB 3 GTTCGACGCTATTGTGAAGAACTCTCGAGTGGTTTCCAGTCAGGCGATCATCGAAGCCGA 62
QY 98 CACGCTCACCGAGGAAGAGCTGTCGCCACCATCGGTACCTGTGGCGCTCGACAGGG 157
DB 63 CATCTCTGTGAAGAGCGCCGGATCGCGCCATCAGGAGAGCCCTCGAGGCCGGAAGC 122
QY 158 CCAGCACACGATGACGCCGCCGCCGCTCGAGGTCCCGGTGAGGTGCGACGA 211
DB 123 CGCCGGACCATCATGATCGCGCAGCAGGTTCGTGATGCCCGGTGTGGTCGATGA 176

RESULT 9
US-09-737-149-3
; Sequence 3, Application US/09737149
; Patent No. US20020077466A1
; GENERAL INFORMATION:
; APPLICANT: Spaderna, Steven K
; APPLICANT: Quinn, Kerry E.
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Muralidhara, Padigaru
; APPLICANT: Spytek, Kimberly A.
; TITLE OF INVENTION: Polypeptides and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-620 CIP
; CURRENT APPLICATION NUMBER: US/09/737,149
; CURRENT FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/170,564
; PRIOR FILING DATE: 1999-12-14
; PRIOR APPLICATION NUMBER: 60/173,165
; PRIOR FILING DATE: 1999-12-27
; PRIOR APPLICATION NUMBER: 60/173,362
; PRIOR FILING DATE: 1999-12-27
; PRIOR APPLICATION NUMBER: 60/173,544
; PRIOR FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: 60/174,404
; PRIOR FILING DATE: 2000-01-04
; PRIOR APPLICATION NUMBER: 60/174,962
; PRIOR FILING DATE: 2000-01-07
; PRIOR APPLICATION NUMBER: 60/223,929
```

```
; PRIOR FILING DATE: 2000-08-09
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2898
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2898)
US-09-737-149-3
```

```
Query Match      17.6%; Score 37.6; DB 10; Length 2898;
Best Local Similarity 48.6%; Pred. No. 0.51;
Matches 103; Conservative 0; Mismatches 109; Indels 0; Gaps 0;

QY 2 CAAGGAGAAAGCGCTACGACCTGGCCGGCTGGCGCGGTACAAGGTGACAAAGAGCTGGG 61
DB 1833 CAAGGGCTTCCGCTTCGGCACCGTGTGGAGAGAGCGCGCGGCGGTACATCAAGAAG 1892
QY 62 TCTTGGCGGTGCCAACCCGGCTCTGTGACTGCCACACGCTCACCGAGGAAGAGCTGT 121
DB 1893 CTTCCCGGACATGCACGCACACATGCGGCGCCACAGCGCGCCACACGCGCGCGGT 1952
QY 122 CGCCACCATCGGTACTCTGTGGCTGCACGAGGCGCCAGACCATGACCGCCCGGG 181
DB 1953 CGCCATGCTCACGAGCGACCCCGCCAAAGCTCAACGCTTTCATCATGGACAGTCGCTCT 2012
QY 182 CGGCGCTCGAGGTCCGCGTTCGAGGTGCGACGACA 213
DB 2013 GGACTAGAGAGGTCTCCATCGACGCGGACTGCA 2044
```

```
RESULT 10
US-09-867-550-669
; Sequence 669, Application US/09867550
; Patent No. US20020082206A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Mehraban, Fuad,
; APPLICANT: Conley, Pamela
; APPLICANT: Law, Debbie
; APPLICANT: Topper, James
; TITLE OF INVENTION: No. US20020082206A1el Polynucleotides from Atherogenic Cells a
; TITLE OF INVENTION: Thereby
; FILE REFERENCE: 21402-013 (Cura-313)
; CURRENT APPLICATION NUMBER: US/09/867,550
; CURRENT FILING DATE: 2001-09-20
; PRIOR APPLICATION NUMBER: USSN 60/208,427
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 2125
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 669
; LENGTH: 584
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (31)..(584)
; OTHER INFORMATION: wherein any n is one of a or t or g or c
US-09-867-550-669
```

```
Query Match      17.1%; Score 36.6; DB 10; Length 584;
Best Local Similarity 53.1%; Pred. No. 0.78;
Matches 78; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

QY 68 CGGTGCCAACCCGGCTCTGTGACTGCCACACGCTCACCGAGGAAGAGCTGTGCGCCAC 127
DB 168 CAGCGGCTCGCGCTGTCCACGATCGCGGTGATGCTGCGCGCGGTGCGGATCATGATGAC 227
QY 128 CATCGGTTACCTGCGCTGCGCTGACGAGGGGCCAGACACGATGACCGCCCGGGGGCT 187
DB 228 CCTCGCGGACCTGAACGGGATCCACCGATCCGATCGACATCTCGATGCCGAGGGGCGTGC 287
```

QY 188 CGAGGTCCTCCGGTTCGAGGTCGACGACAT 214  
 Db 288 GCAGAACACCGGGCGAGGTCACATCCT 314

RESULT 11

```

US-09-864-761-26660
: Sequence 26660, Application US/09864761
: Patent NO. US20020048763A1
: GENERAL INFORMATION:
: APPLICANT: Penn, Sharron G.
: APPLICANT: Rank, David R.
: APPLICANT: Hanzel, David K.
: APPLICANT: Chen, Wensheng
: TITLE OF INVENTION: HUMAN GENOME-DERIVED
: TITLE OF INVENTION: GENE EXPRESSION ANAL
: FILE REFERENCE: Aecmica-X-1
: CURRENT APPLICATION NUMBER: US/09/864,761
: CURRENT FILING DATE: 2001-05-23
: PRIOR APPLICATION NUMBER: US 60/180,312
: PRIOR FILING DATE: 2000-02-04
: PRIOR APPLICATION NUMBER: US 60/207,456
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: US 09/632,366
: PRIOR FILING DATE: 2000-08-03
: PRIOR APPLICATION NUMBER: GB 24263.6
: PRIOR FILING DATE: 2000-10-04
: PRIOR APPLICATION NUMBER: US 60/236,359
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: PCT/US01/00666
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00667
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00664
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00669
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00665
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00668
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00663
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00662
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00661
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00670
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: US 60/234,687
: PRIOR FILING DATE: 2000-09-21
: PRIOR APPLICATION NUMBER: US 09/608,408
: PRIOR FILING DATE: 2000-06-30
: PRIOR APPLICATION NUMBER: US 09/774,203
: PRIOR FILING DATE: 2001-01-29
: NUMBER OF SEQ ID NOS: 49117
: SOFTWARE: Anomax Sequence Listing Engine
: SEQ ID NO 26660
: LENGTH: 221

```

LENGTH: 221  
TYPE: DNA

```

1 TYPE: DNA
2 ORGANISM: Homo sapiens
3 FEATURE:
4 OTHER INFORMATION: MAP TO AL049569.13
5 OTHER INFORMATION: EXPRESSED IN BT474. SIGNAL = 1.3
6 OTHER INFORMATION: EXPRESSED IN BONE MARROW. SIGNAL = 1.9
7 OTHER INFORMATION: EXPRESSED IN BRAIN. SIGNAL = 1
8 OTHER INFORMATION: NT HIT: AB030176.1, EVALUAE 4.00e-97
9 OTHER INFORMATION: EST_HUMAN HIT: BE089551.1, EVALUAE 1.00e-94
10 OTHER INFORMATION: SWISSPROT HIT: Q9YJ28, EVALUAE 1.00e-29
11 US-09-864-761-26660

```

Query Match	16.9%	Score 36.2	DB 10	Length 221
-------------	-------	------------	-------	------------

Best Local Similarity 50.3%; Pred. No. 0.91;  
Matches 89; Conservative 0; Mismatches

Qy	14	CTACGACCTGCCCGCGTGGGCCGGTACAAAGTGAACAAGAAGCTTGGCTCTTGGCGGTGC	73
Db	26	CCACATCCTGGCGCGGGAAGCTTACCATGTGTCAGTACACGGTGGCTCCGCGGA	85
Qy	74	CAACCCGCTCTGCTGACTGCCACACAGCTCACCGAGAAGAGTCTCTGCCACACCTCG	133
Db	86	GCTGCTGTTCTCTGGAAGGCTCTGTTCCCGACGAGGCTTCTCAGGCTTGGTCTC	145
Qy	134	GTACCTGGTGGCCTGTGACGAGGCGCAGACCATGACCGCCCCCGGCGGCTCGA	190
Db	146	CATCATGTTCAGCCTGTGTGAGTACATGGCCAGGTGAGGCTCCAGGCTCCACCA	202

## RESULT 12

```

US-09-878-574-3598
; Sequence 3598, Application US/09878574
; Patent No. US20020110548A1
; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Thompson, Michael D.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(15401)B
; CURRENT APPLICATION NUMBER: US/09/878,574
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 09/333,535
; PRIOR FILING DATE: 1999-06-14
; NUMBER OF SEQ ID NOS: 15775
; SEQ ID NO 3598
; LENGTH: 392
; TYPE: DNA
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: LIB3028-008-Q1-B1-C6
US-09-878-574-3598

```

### Query Match

Query Match	16.9%	Score 36.2	DB 10	Length 392	
Best Local Similarity	54.0%	Pred. No. 0.95			
Matches	74	Conservative 0	Mismatches 63	Indels 0	Gaps 0
QY	76	ACCCGGCTCTGTGACTGCCACCAAGCTACCGAGGAGACGTCGTGCGCCACCATCGGT	135		
Db	2	ACCTGCCCCAGGCGCTTCGGCCTCATCGACCGCGACACGACGGCGTCTCACGCGCCAGG	61		
QY	136	ACCTGTGTGGCCTTCACACGAGGGCCAGCACCGATGACGCCGCCCGCGGCGCTCGAGGTCC	195		
Db	62	ACCTCGAGSCCCTCTCACTGCTCTCGCGCGTGCCTGCCCGAGACGCTGGCGGTCA	121		
QY	196	CGGTGAGGTTCAGAC	212		
Db	122	TGTCGGCGAGGTCGAC	138		

RESULT 13

```

US-09-815-242-7960
; Patent 7960, Application US/09815242
; Sequence No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Ess
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242

```

```

; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7960
; LENGTH: 1914
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; NAME/KEY: CDS
; LOCATION: (1)...(1914)
US-09-815-242-7960

Query Match 16.9%; Score 36.2; DB 10; Length 1914;
Best Local Similarity 53.1%; Pred. No. 1.1;
Matches 77; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 69 GGTGCCAACCCGGCTGTGGTACTGCCACACGCTCACCGAGGAAGACGTCTGTCGCCACC 128
   || || || || || || || || || || || || || || || || || || || || ||
Db 1615 GCGACGCGGTGTCCACGCGACCGCGCAAGATGATCACCGAGCGCGGACAAAGCCACC 1674

QY 129 ATCGGTACTGTGTGGCTGTCAGAGGGCCAGACACCATGACCGCCCGCGGCGCTC 188
   || || || || || || || || || || || || || || || || || || || || ||
Db 1675 GCCGAGCAAGCGGACCATCGAAGAGCGCTGGCGGAGCTGGAAGCGCGGTGAAGGCC 1734

QY 189 GAGTCTCCCGTCGAGTTCGACGACA 213
   || || || || || || || || || ||
Db 1735 GACGACAAGCGCGAGATCGAGGCCA 1759

RESULT 14
US-09-815-242-4131
; Sequence 4131, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7960
; LENGTH: 1914
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; NAME/KEY: CDS
; LOCATION: (1)...(1914)
US-09-815-242-7960

Query Match 16.9%; Score 36.2; DB 10; Length 1914;
Best Local Similarity 53.1%; Pred. No. 1.1;
Matches 77; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 69 GGTGCCAACCCGGCTGTGGTACTGCCACACGCTCACCGAGGAAGACGTCTGTCGCCACC 128
   || || || || || || || || || || || || || || || || || || || || ||
Db 1615 GCGACGCGGTGTCCACGCGACCGCGCAAGATGATCACCGAGCGCGGACAAAGCCACC 1674

QY 129 ATCGGTACTGTGTGGCTGTCAGAGGGCCAGACACCATGACCGCCCGCGGCGCTC 188
   || || || || || || || || || || || || || || || || || || || || ||
Db 1675 GCCGAGCAAGCGGACCATCGAAGAGCGCTGGCGGAGCTGGAAGCGCGGTGAAGGCC 1734

QY 189 GAGTCTCCCGTCGAGTTCGACGACA 213
   || || || || || || || || || ||
Db 1735 GACGACAAGCGCGAGATCGAGGCCA 1759

RESULT 15
US-09-923-876-2475
; Sequence 2475, Application US/09923876
; Patent No. US20030013958A1
; GENERAL INFORMATION:
; APPLICANT: Laigudi, Raghunath V.
; APPLICANT: Kamigaki, Laura Y. (Ito)
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN SEEDLING
; FILE REFERENCE: PL-0015-1 CON
; CURRENT APPLICATION NUMBER: US/09/923,876
; CURRENT FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: 09/298,329
; PRIOR FILING DATE: 1999-04-21
; PRIOR APPLICATION NUMBER: 60/085,331
; PRIOR FILING DATE: 1998-05-05
; NUMBER OF SEQ ID NOS: 6332
; SOFTWARE: PERL Program
; SEQ ID NO 2475
; LENGTH: 242
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc-feature
; OTHER INFORMATION: Incyte ID No. US20020013958A1 700160665H1
US-09-923-876-2475

Query Match 16.6%; Score 35.6; DB 10; Length 242;
Best Local Similarity 57.0%; Pred. No. 1.3;
Matches 65; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 99 AGCTCTACCGAGGAAGACGTGTCGCCACCATCGGTACCTGGTGCCTGCACGAGGC 158
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 124 ATGTACCTCTGACGTGGAGACCGCCGCCACAGCAGCGTCAAGACCCACGGTCTCGCCCG 183

QY 159 CAGACGACGATGACCGCCCGCGGCGCTCCAGGTCGGGTCCGGTTCGAGGTCGAGGAC 212
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 184 CCGACCAACAGGCCAGCGCGCGGCTTCGAGGCGCCACCGAGGCGCGCGC 237

Search completed: November 12, 2002, 16:57:53
Job time : 38.7711 secs
```

```

; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4131
; LENGTH: 3057
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-815-242-4131

Query Match 16.9%; Score 36.2; DB 10; Length 3057;
Best Local Similarity 55.0%; Pred. No. 1.1;
Matches 71; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 74 CAACCCGGCTCTGTGACTGCCACACGCTCACCGAGGAAGACGTCTGTCGCCACCATCGG 133
   || || || || || || || || || || || || || || || || || || || || ||
Db 483 CCAGCCCAAGCTGCGGACCTGCCCGGTATCGCCGAGGCGGAGATCTCGGCAACCAAGT 542

QY 134 GTACCTGGTGGCTGACGAGGCGCAGACCATGATGACCGCCCGCGGCGCTCGAGGT 193
   || || || || || || || || || || || || || || || || || || || || ||
Db 543 GTTCGCCATCGCGCTGTGCTGGACCCGCTGAAGATGGCGCGTTCGCGCTCACCGCGG 602

QY 194 CCGGCTCGA 202
   | | | | |
Db 603 CGAGATCAA 611

RESULT 15
US-09-923-876-2475
; Sequence 2475, Application US/09923876
; Patent No. US20030013958A1
; GENERAL INFORMATION:
; APPLICANT: Laigudi, Raghunath V.
; APPLICANT: Kamigaki, Laura Y. (Ito)
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN SEEDLING
; FILE REFERENCE: PL-0015-1 CON
; CURRENT APPLICATION NUMBER: US/09/923,876
; CURRENT FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: 09/298,329
; PRIOR FILING DATE: 1999-04-21
; PRIOR APPLICATION NUMBER: 60/085,331
; PRIOR FILING DATE: 1998-05-05
; NUMBER OF SEQ ID NOS: 6332
; SOFTWARE: PERL Program
; SEQ ID NO 2475
; LENGTH: 242
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc-feature
; OTHER INFORMATION: Incyte ID No. US20020013958A1 700160665H1
US-09-923-876-2475

Query Match 16.6%; Score 35.6; DB 10; Length 242;
Best Local Similarity 57.0%; Pred. No. 1.3;
Matches 65; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 99 AGCTCTACCGAGGAAGACGTGTCGCCACCATCGGTACCTGGTGCCTGCACGAGGC 158
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 124 ATGTACCTCTGACGTGGAGACCGCCGCCACAGCAGCGTCAAGACCCACGGTCTCGCCCG 183

QY 159 CAGACGACGATGACCGCCCGCGGCGCTCCAGGTCGGGTCCGGTTCGAGGTCGAGGAC 212
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 184 CCGACCAACAGGCCAGCGCGCGGCTTCGAGGCGCCACCGAGGCGCGCGC 237

Search completed: November 12, 2002, 16:57:53
Job time : 38.7711 secs
```



Result No.	Score	Query Match	Length	DB	ID	Description	
1	124.8	58.3	970	1	US-08-250-030-1	Sequence 1, Appl	
2	124.8	58.3	970	5	PTC-US95-06790-1	Sequence 1, Appl	
3	104	48.6	3447	2	US-08-313-185-57	Sequence 57, Appl	
4	104	48.6	3447	3	US-09-082-61A-57	Sequence 1, Appl	
5	43.2	20.2	4411529	4	US-09-103-840A-1	Sequence 10, Appl	
6	41.2	19.3	1620	2	US-08-461-775-10	Sequence 10, Appl	
7	41.2	19.3	1620	3	US-09-031-606-10	Sequence 10, Appl	
8	41.2	19.3	2668	2	US-08-461-775-11	Sequence 11, Appl	
9	41.2	19.3	2668	3	US-09-031-606-11	Sequence 11, Appl	
10	41	19.2	1926	4	US-09-073-985-10	Sequence 10, Appl	
11	41	19.2	3000	4	US-09-192-104-1	Sequence 1, Appl	
12	41	19.2	3000	4	US-09-543-446-1	Sequence 1, Appl	
13	40	18.7	1665	3	US-08-881-784-8	Sequence 8, Appl	
14	40	18.7	1665	4	US-09-292-768-3	Sequence 3, Appl	
15	40	18.7	1665	4	US-09-292-768-67	Sequence 67, Appl	
16	40	18.7	1665	4	US-09-292-768-69	Sequence 69, Appl	
17	40	18.7	15872	4	US-09-105-537-1	Sequence 1, Appl	
18	39	18.2	1734	6	5352575-8	Patent No. 5352575	
19	36.8	17.2	1950	3	US-08-911-853-14	Sequence 14, Appl	
20	36.8	17.2	1950	4	US-09-479-409-14	Sequence 14, Appl	
21	36.8	17.2	1950	4	US-09-479-453-14	Sequence 14, Appl	
22	36.8	17.2	17612	3	US-08-911-853-29	Sequence 29, Appl	
23	36.8	17.2	17612	4	US-09-479-409-29	Sequence 29, Appl	
24	36.8	17.2	17612	4	US-09-479-453-29	Sequence 29, Appl	
25	36.8	17.2	68750	3	US-09-335-409-1	Sequence 1, Appl	
26	36.8	17.2	68750	4	US-09-568-102-1	Sequence 1, Appl	
27	36.8	17.2	68750	4	US-09-567-969-1	Sequence 1, Appl	





```
; NAME: Crane-Feury, Sharon E
; REGISTRATION NUMBER: 36,113
; REFERENCE/DOCKET NUMBER: 010830-035
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1620 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1620
; US-09-031-606-10

Query Match          19.3%; Score 41.2; DB 2; Length 1620;
Best Local Similarity 49.5%; Pred. No. 0.12;
Matches 106; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

QY 1 TCAAGGAGAACCGCTACGACCTGGCCCGCTGGCGGTACAAGGTGAACAAGAAGCTGG 60
    |||||  |  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 1268 TCAAGGTCTCTGGAGCACAACCTCGCGCCACCGCGCAGCAGGACCGCGGTGCGGGTGG 1327
    |||||  |  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

QY 61 GTCTTGGCGGTGCCAACCCGGCTCTGTGACTGCCACCGCTCACCGAGGAGAGAGCTGG 120
    |||||  |  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 1328 TCCGCGCGCGCGCTCGAGCCGCTCGGTGGATCGCCGAGAACCGCGCCCTCGAGGGCT 1387
    |||||  |  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

QY 121 TCGCCACCATCGGTACTCGTGGCTGCACGAGGCGCCAGACGATGACCGCCCGCG 180
    |||||  |  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 1388 AGTCATCACCAACAGGTGGCGAGCTGACAAGGCGCGGGCTTCAACGCGGCCACCG 1447
    |||||  |  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

QY 181 GCGGCTTCGAGGTCCGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 214
    |||||  |  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 1448 GCGAGTACGGCGCTGGTCAAGCGCGCGGTCTCAT 1481

RESULT 7
US-09-031-606-10
; Sequence 10, Application US/09031606
; Patent No. 6153404
; GENERAL INFORMATION:
; APPLICANT: MAZODIER, Philippe
; APPLICANT: GUGLIEMI, Gerard
; TITLE OF INVENTION: REGULATORY NUCLEOTIDE SEQUENCE OF THE
; TITLE OF INVENTION: INITIATION OF TRANSCRIPTION
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: George Mason Bldg., Washington & Prince Sts.
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/031,606
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/050,313
; FILING DATE: 10-MAY-1993
; APPLICATION NUMBER: FR 9011186
; FILING DATE: 10-SEP-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Crane-Feury, Sharon E
; REGISTRATION NUMBER: 36,113
; TELECOMMUNICATION INFORMATION:

; NAME: Crane-Feury, Sharon E
; REGISTRATION NUMBER: 36,113
; REFERENCE/DOCKET NUMBER: 010830-035
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1620 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1620
; US-08-461-775-11

Query Match          19.3%; Score 41.2; DB 3; Length 1620;
Best Local Similarity 49.5%; Pred. No. 0.12;
Matches 106; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

QY 1 TCAAGGAGAACCGCTACGACCTGGCCCGCTGGCGGTACAAGGTGAACAAGAAGCTGG 60
    |||||  |  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 1368 TCAAGGTCTCTGGAGCACAACCTCGCGCCACCGCGCAGCAGGACCGCGGTGCGGGTGG 1327
    |||||  |  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

QY 61 GTCTTGGCGGTGCCAACCCGGCTCTGTGACTGCCACCGCTCACCGAGGAGAGAGCTGG 120
    |||||  |  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 1328 TCCGCGCGCGCGCTCGAGCCGCTCGGTGGATCGCCGAGAACCGCGCCCTCGAGGGCT 1387
    |||||  |  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

QY 121 TCGCCACCATCGGTACTCGTGGCTGCACGAGGCGCCAGACGATGACCGCCCGCG 180
    |||||  |  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 1388 AGTCATCACCAACAGGTGGCGAGCTGACAAGGCGCGGGCTTCAACGCGGCCACCG 1447
    |||||  |  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

QY 181 GCGGCTTCGAGGTCCGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 214
    |||||  |  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 1448 GCGAGTACGGCGCTGGTCAAGCGCGCGGTCTCAT 1481

RESULT 8
US-08-461-775-11
; Sequence 11, Application US/08461775
; Patent No. 5858773
; GENERAL INFORMATION:
; APPLICANT: MAZODIER, Philippe
; APPLICANT: GUGLIEMI, Gerard
; TITLE OF INVENTION: REGULATORY NUCLEOTIDE SEQUENCE OF THE
; TITLE OF INVENTION: INITIATION OF TRANSCRIPTION
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: George Mason Bldg., Washington & Prince Sts.
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/461,775
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/050,313
; FILING DATE: 10-MAY-1993
; APPLICATION NUMBER: FR 9011186
; FILING DATE: 10-SEP-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Crane-Feury, Sharon E
; REGISTRATION NUMBER: 36,113
; REFERENCE/DOCKET NUMBER: 010830-035
; TELECOMMUNICATION INFORMATION:
```

TELEPHONE: (703) 836-6620  
TELEFAX: (703) 836-2021  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2668 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-461-775-11

Query Match 19.3%; Score 41.2; DB 2; Length 2668;  
Best Local Similarity 49.5%; Pred. No. 0.13;  
Matches 106; Conservative 0; Mismatches 108; Indels 0; Gaps 0;  
QY 1 TCAAGGAGAAGCGCTACGACCTGGCCCGCTGGCCCGGTACAAGGTGAACAAGAGCTGG 60  
Db 2116 TCAAGGTCTTGACGACAACTTCGGCCGACCGCGGACGAGGCCACCGGTGTGCGCGTCG 2175  
QY 61 GTCTTGGCGGTGCCAACCCGGCTCTGGTGACTGCCACACCGCTCACCGAGGAAGACGTGG 120  
Db 2176 TCCGCGCGCGCGCTCGAGCGCTGGCTGGATCGCGGAGAACGCCGCGCTCGAGGGCT 2235  
QY 121 TCGCCACCATCGGTACCTGGTGGCCCTGACGAGGGCCAGACACGATACCGCCCCCG 180  
Db 2236 ACGTATCACCACCAAGGTGGCGAGCTCGACAAGGGCCAGGGCTTCAACCGCGCCACCG 2295  
QY 181 CGCGGCTCGAGTCCCGCTCGAGTTCGACGACAT 214  
Db 2296 GCGAGTACGGGACCTGGTCAAGCGCGCGTCAT 2329

RESULT 9  
US-09-031-606-11  
Sequence 11, Application US/09031606  
Patent No. 6153404  
GENERAL INFORMATION:  
APPLICANT: MAZODIER, Philippe  
APPLICANT: GUGLIEMI, Gerard  
TITLE OF INVENTION: REGULATORY NUCLEOTIDE SEQUENCE OF THE  
INITIATION OF TRANSCRIPTION  
NUMBER OF SEQUENCES: 15  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Burns, Doane, Swecker & Mathis  
STREET: George Mason Bldg., Washington & Prince Sts.  
CITY: Alexandria  
STATE: Virginia  
COUNTRY: United States  
ZIP: 22313-1404  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/031,606  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/050,313  
FILING DATE: 10-MAY-1993  
APPLICATION NUMBER: FR 9011186  
FILING DATE: 10-SEP-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Crane-Feury, Sharon E  
REGISTRATION NUMBER: 36,113  
REFERENCE/DOCKET NUMBER: 010830-035  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 836-6620  
TELEFAX: (703) 836-2021  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2668 base pairs

TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-09-031-606-11

Query Match 19.3%; Score 41.2; DB 3; Length 2668;  
Best Local Similarity 49.5%; Pred. No. 0.13;  
Matches 106; Conservative 0; Mismatches 108; Indels 0; Gaps 0;  
QY 1 TCAAGGAGAAGCGCTACGACCTGGCCCGCTGGCCCGGTACAAGGTGAACAAGAGCTGG 60  
Db 2116 TCAAGGTCTTGACGACAACTTCGGCCGACCGCGGACGAGGCCACCGGTGTGCGCGTCG 2175  
QY 61 GTCTTGGCGGTGCCAACCCGGCTCTGGTGACTGCCACACCGCTCACCGAGGAAGACGTGG 120  
Db 2176 TCCGCGCGCGCGCTCGAGCGCTGGCTGGATCGCGGAGAACGCCGCGCTCGAGGGCT 2235  
QY 121 TCGCCACCATCGGTACCTGGTGGCCCTGACGAGGGCCAGACACGATACCGCCCCCG 180  
Db 2236 ACGTATCACCACCAAGGTGGCGAGCTCGACAAGGGCCAGGGCTTCAACCGCGCCACCG 2295  
QY 181 CGCGGCTCGAGTCCCGCTCGAGTTCGACGACAT 214  
Db 2296 GCGAGTACGGGACCTGGTCAAGCGCGCGTCAT 2329

RESULT 10  
US-09-079-955-10  
Sequence 10, Application US/09079955A  
Patent No. 6465209  
GENERAL INFORMATION:  
APPLICANT: Alexander Blinkovsky  
APPLICANT: Kimberly Brown  
APPLICANT: Elizabeth Gollightly  
APPLICANT: Tony Byun  
APPLICANT: Thomas Mathiasen  
APPLICANT: Lene V. Kofod  
APPLICANT: Mikio Fujii  
APPLICANT: Chigusa Shizuoka  
TITLE OF INVENTION: Methods For Producing Protein  
FILE REFERENCE: 5253 500-US  
CURRENT APPLICATION NUMBER: US/09/079,955A  
CURRENT FILING DATE: 1998-05-15  
NUMBER OF SEQ ID NOS: 11  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 10  
LENGTH: 1926  
TYPE: DNA  
ORGANISM: Sphingomonas capsulata  
US-09-079-955-10

Query Match 19.2%; Score 41; DB 4; Length 1926;  
Best Local Similarity 51.4%; Pred. No. 0.14;  
Matches 95; Conservative 0; Mismatches 90; Indels 0; Gaps 0;  
QY 27 CGCGTGGCGCGGTACAAGGTGAACAAGAGCTGGGTCTTTGGCGGTGCCAACCCGGCTCTG 86  
Db 1132 CGCGACTGGCGCTCGGTGGAAGACACGACGATGATCCCATCTTCGCGCGCCGCAAGCCC 1191  
QY 87 GTGACTGCCACCGCTCACCGAGGAAGAGCGTCTGCCACCATCGGGTACCTGGTGGCG 146  
Db 1192 AAGCCCTATTCTTCGCTTACCCGTAACGAGGACTATTACACCGAAGGCGCGCTGGTGTG 1251  
QY 147 CTGACGAGGGCCAGCACGATGACCGCCCCCGCGGCTCGAGGTCCCGGTCTCGAGGTC 206  
Db 1252 CTGGAAGCGGACCAAGATCATCCCGGATGGCACCGGGCGCAAGAGGGCTTGATGATTC 1311  
QY 207 GACCA 211  
Db 1312 GCCAA 1316

```
RESULT 11
US-09-192-104-1
; Sequence 1, Application US/09192104B
; Patent No. 6184020
; GENERAL INFORMATION:
; APPLICANT: Alexander Blinkovsky
; APPLICANT: Tony Byun
; APPLICANT: Alan V. Klotz
; APPLICANT: Alan Sloma
; APPLICANT: Maria Tang
; APPLICANT: Mikio Fujii
; APPLICANT: Chigusa Marumoto
; APPLICANT: Lene Venke Kofod
; TITLE OF INVENTION: Polypeptides Having Amino-peptidase
; FILE REFERENCE: 5379.200-US
; CURRENT APPLICATION NUMBER: US/09/192,104B
; CURRENT FILING DATE: 1998-11-13
; EARLIER APPLICATION NUMBER: 60/069719
; EARLIER FILING DATE: 1997-12-16
; EARLIER APPLICATION NUMBER: 1465/97
; EARLIER FILING DATE: 1997-12-16
; EARLIER APPLICATION NUMBER: PA 1998 00670
; EARLIER FILING DATE: 1998-05-15
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3000
; TYPE: DNA
; ORGANISM: Sphingomonas
US-09-192-104-1

Query Match      19.2%; Score 41; DB 4; Length 3000;
Best Local Similarity 51.4%; Pred. No. 0.14;
Matches 95; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

QY 27 CGCGTGGCGCGGTACAAAGCTGAACAGAGCTGGTCTTTGGCGGTGCCAACCCGGCTCTG 86
    |||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1801 CGCGACTGGCGCTCGGTGGAAGACACGACGATGGATCCATCTTCGCCGCCGCAAGCCC 1860

QY 87 GTGACTGCCACCAAGCTCAGCGAGGAGAGCTGCTGCCACCATCGGGTACCTGGTGGC 146
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1861 AAGCCCTATTCTCGCTTACCGGTAAACGAGGACTATTACACCGAAGCGCGCTGGTGG 1920

QY 147 CTGCGAGGCGGACACCATGATCGCGCGCCCGCGGCTCGAGGTCCCGGTGAGGTC 206
    |||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1921 CTGGAAGCGGACCATGATCGCGGATGGCGGCAAGAGCGGCTGGATGATTC 1980

QY 207 GACGA 211
    | | |
Db 1981 GCCAA 1985

RESULT 12
US-09-543-446-1
; Sequence 1, Application US/09543446
; Patent No. 6303360
; GENERAL INFORMATION:
; APPLICANT: Alexander Blinkovsky
; APPLICANT: Tony Byun
; APPLICANT: Alan V. Klotz
; APPLICANT: Alan Sloma
; APPLICANT: Maria Tang
; APPLICANT: Mikio Fujii
; APPLICANT: Chigusa Marumoto
; APPLICANT: Lene Venke Kofod
; TITLE OF INVENTION: Polypeptides Having Amino-peptidase
; FILE REFERENCE: 5379.210-US
; CURRENT APPLICATION NUMBER: US/09/543,446
; CURRENT FILING DATE: 2000-04-05
; EARLIER APPLICATION NUMBER: 60/069719
; EARLIER FILING DATE: 1997-12-16
; EARLIER APPLICATION NUMBER: 1465/97
; EARLIER FILING DATE: 1997-12-16
; EARLIER APPLICATION NUMBER: PA 1998 00670
; EARLIER FILING DATE: 1998-05-15
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3000
; TYPE: DNA
; ORGANISM: Sphingomonas
US-09-192-104-1

Query Match      19.2%; Score 41; DB 4; Length 3000;
Best Local Similarity 51.4%; Pred. No. 0.14;
Matches 95; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

QY 27 CGCGTGGCGCGGTACAAAGCTGAACAGAGCTGGTCTTTGGCGGTGCCAACCCGGCTCTG 86
    |||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1801 CGCGACTGGCGCTCGGTGGAAGACACGACGATGGATCCATCTTCGCCGCCGCAAGCCC 1860

QY 87 GTGACTGCCACCAAGCTCAGCGAGGAGAGCTGCTGCCACCATCGGGTACCTGGTGGC 146
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1861 AAGCCCTATTCTCGCTTACCGGTAAACGAGGACTATTACACCGAAGCGCGCTGGTGG 1920

QY 147 CTGCGAGGCGGACACCATGATCGCGCGCCCGCGGCTCGAGGTCCCGGTGAGGTC 206
    |||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1921 CTGGAAGCGGACCATGATCGCGGATGGCGGCAAGAGCGGCTGGATGATTC 1980

QY 207 GACGA 211
    | | |
Db 1981 GCCAA 1985

RESULT 13
US-08-881-784-8
; Sequence 8, Application US/08881784
; Patent No. 6083731
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney B.
; APPLICANT: Lupien, Shari L.
; APPLICANT: Karp, Frank
; TITLE OF INVENTION: RECOMBINANT MATERIALS AND METHODS FOR
; TITLE OF INVENTION: THE PRODUCTION OF LIMONENE HYDROXYLASES
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Christensen, O'Connor, Johnson and Kindness
; ADDRESS: PLLC
; STREET: 1420 Fifth Avenue, Suite 2800
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/881,784
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Shelton, Dennis K.
; REGISTRATION NUMBER: 26,997
; REFERENCE/DOCKET NUMBER: WSUR19777
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 224-0718
; TELEFAX: (206) 224-0779
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
```

```
; EARLIER FILING DATE: 1997-12-16
; EARLIER APPLICATION NUMBER: 1465/97
; EARLIER FILING DATE: 1997-12-16
; EARLIER APPLICATION NUMBER: PA 1998 00670
; EARLIER FILING DATE: 1998-05-15
; EARLIER APPLICATION NUMBER: 09/192,104
; EARLIER FILING DATE: 1998-11-13
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3000
; TYPE: DNA
; ORGANISM: Sphingomonas
US-09-543-446-1

Query Match      19.2%; Score 41; DB 4; Length 3000;
Best Local Similarity 51.4%; Pred. No. 0.14;
Matches 95; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

QY 27 CGCGTGGCGCGGTACAAAGCTGAACAGAGCTGGTCTTTGGCGGTGCCAACCCGGCTCTG 86
    |||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1801 CGCGACTGGCGCTCGGTGGAAGACACGACGATGGATCCATCTTCGCCGCCGCAAGCCC 1860

QY 87 GTGACTGCCACCAAGCTCAGCGAGGAGAGCTGCTGCCACCATCGGGTACCTGGTGGC 146
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1861 AAGCCCTATTCTCGCTTACCGGTAAACGAGGACTATTACACCGAAGCGCGCTGGTGG 1920

QY 147 CTGCGAGGCGGACACCATGATCGCGCGCCCGCGGCTCGAGGTCCCGGTGAGGTC 206
    |||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1921 CTGGAAGCGGACCATGATCGCGGATGGCGGCAAGAGCGGCTGGATGATTC 1980

QY 207 GACGA 211
    | | |
Db 1981 GCCAA 1985

RESULT 13
US-08-881-784-8
; Sequence 8, Application US/08881784
; Patent No. 6083731
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney B.
; APPLICANT: Lupien, Shari L.
; APPLICANT: Karp, Frank
; TITLE OF INVENTION: RECOMBINANT MATERIALS AND METHODS FOR
; TITLE OF INVENTION: THE PRODUCTION OF LIMONENE HYDROXYLASES
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Christensen, O'Connor, Johnson and Kindness
; ADDRESS: PLLC
; STREET: 1420 Fifth Avenue, Suite 2800
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/881,784
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Shelton, Dennis K.
; REGISTRATION NUMBER: 26,997
; REFERENCE/DOCKET NUMBER: WSUR19777
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 224-0718
; TELEFAX: (206) 224-0779
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
```

```
; LENGTH: 1665 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; ORIGINAL SOURCE:
; ORGANISM: Mentha x piperita
; IMMEDIATE SOURCE:
; CLONE: pPM17
US-08-881-784-8

Query Match      18.7%; Score 40; DB 3; Length 1665;
Best Local Similarity 50.5%; Pred. No. 0.23;
Matches 97; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

QY 23 GCGCCCGGTGGCGGTGACAAAGCTGTAACAAAGCTGGTCTTGGCGGTGCCAACCCGGC 82
Db 270 GTCCCGGAGGCCACGAAGAGGCGATGAAGCTGTGGACCGCGCTCGCGGACCGGTT 329

QY 83 TCTGGTGACTGCCACCGCTCACCGAGGAGAGCTGTCGCCACCATCGGTACTCGT 142
Db 330 CGAGAGCATCGGACGAAGATCATGTGGTACGACACGACATCATCTTCAGCCCTTA 389

QY 143 GCGCCTGCACGAGGCCACGACCATGACCGCCCGCGGCTCGAGGTCCCGGTGCA 202
Db 390 CAGCGTCACTGGCGCCAGATCGGGAAGATCTCGCTCTCCGAGCTCTCAGCGCCGCA 449

QY 203 GGTCCGACGACAT 214
Db 450 CGTCCGCTCCTT 461

RESULT 14
US-09-292-768-3
; Sequence 3, Application US/09292768
; Patent No. 6194185
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney B
; APPLICANT: Lupien, Shari L
; APPLICANT: Karp, Frank
; TITLE OF INVENTION: RECOMBINANT MATERIALS AND METHODS FOR THE PRODUCTION OF
; FILE REFERENCE: wsur13463
; CURRENT APPLICATION NUMBER: US/09/292,768
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 08/881,784
; EARLIER FILING DATE: 1997-06-24
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1665
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: computer-generated nucleic acid sequence encoding
; OTHER INFORMATION: limonene-3-hydroxylase
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1665)
; OTHER INFORMATION: computer-generated nucleic acid sequence encoding
; OTHER INFORMATION: limonene-3-hydroxylase variant
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (19)..(1518)
US-09-292-768-67

Query Match      18.7%; Score 40; DB 4; Length 1665;
Best Local Similarity 50.5%; Pred. No. 0.23;
Matches 97; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

QY 23 GCGCCCGGTGGCGGTGACAAAGCTGTAACAAAGCTGGTCTTGGCGGTGCCAACCCGGC 82
Db 270 GTCCCGGAGGCCACGAAGAGGCGATGAAGCTGTGGACCGCGCTCGCGGACCGGTT 329

QY 83 TCTGGTGACTGCCACCGCTCACCGAGGAGAGCTGTCGCCACCATCGGTACTCGT 142
Db 330 CGAGAGCATCGGACGAAGATCATGTGGTACGACACGACATCATCTTCAGCCCTTA 389

QY 143 GCGCCTGCACGAGGCCACGACCATGACCGCCCGCGGCTCGAGGTCCCGGTGCA 202
Db 390 CAGCGTCACTGGCGCCAGATCGGGAAGATCTCGCTCTCCGAGCTCTCAGCGCCGCA 449

QY 203 GGTCCGACGACAT 214
Db 450 CGTCCGCTCCTT 461

Search completed: November 12, 2002, 18:45:22
Job time : 2221.7 secs
```

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 12, 2002, 15:44:35 ; Search time 140.151 Seconds  
(without alignments)  
3438.621 Million cell updates/sec

Title: US-09-697-123B-7  
Perfect score: 214  
Sequence: 1 tcaagagaagcgctacgac.....cgggtcaggctgcagacat 214

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_101002.\*  
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*  
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*  
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*  
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*  
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*  
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*  
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*  
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*  
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*  
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*  
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*  
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*  
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*  
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*  
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*  
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*  
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*  
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*  
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*  
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*  
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*  
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*  
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*  
24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	214	100.0	214	22	AAS05207
2	186.8	87.3	214	22	AAS05221
3	153.6	71.8	208	22	AAS05219
4	146.8	68.6	214	22	AAS05214
5	142.4	66.5	208	22	AAS05215
6	140.4	65.6	211	22	AAS05218
7	139.2	65.0	208	22	AAS05201
8	139.2	65.0	208	22	AAS05211
9	137.6	64.3	208	22	AAS05210

10	136	63.6	208	22	AAS05202	Mycobacterium gord
11	134.4	62.8	208	22	AAS05217	Mycobacterium cels
12	132.8	62.1	208	22	AAS05203	Mycobacterium gord
13	132.8	62.1	208	22	AAS05208	Mycobacterium kans
14	128	59.8	208	22	AAS05224	Mycobacterium xeno
15	124.8	58.3	208	22	AAS05205	Mycobacterium tube
16	124.8	58.3	208	22	AAS05206	Mycobacterium terr
17	124.8	58.3	208	22	AAS05216	Mycobacterium bovi
18	124.8	58.3	970	17	AAT09676	Mycobacterium tube
19	124.8	58.3	3519	22	AAH51976	Mycobacterium tube
20	124.8	58.3	3534	22	AAH02079	Mycobacterium tube
21	124.8	58.3	3853	21	AAA74651	Mycobacterium tube
22	124.8	58.3	3853	21	AAA89994	M. tuberculosis rp
23	123.2	57.6	208	22	AAS05222	Mycobacterium affi
24	120.8	56.4	207	22	AAS05212	Mycobacterium szul
25	120	56.1	205	22	AAS05220	Mycobacterium intr
26	118.4	55.3	208	22	AAS05223	Mycobacterium haem
27	116	54.2	207	22	AAS05204	Mycobacterium gord
28	106.2	49.6	223	22	AAS05209	Mycobacterium scro
29	104	48.6	3447	14	AAQ51532	M. leprae rpoB gene
30	99.8	46.6	223	22	AAS05213	Mycobacterium gast
31	88.4	41.3	3495	22	AAH65512	C glutamicum codin
32	88.4	41.3	34980	22	AAH68525	Propionibacterium
33	56.8	26.5	27426	23	AAS59541	Mycobacterium tube
34	43.2	20.2	9210	22	AAH52046	Human ORFX polynuc
35	43.2	20.2	4411529	22	AAI99682	Pseudorabies virus
36	42.2	19.7	318	24	ABN20109	Maize nitrite redu
37	42	19.6	5895	23	ABN87076	groEL-1 gene. Str
38	41.6	19.4	1844	12	AAQ12000	Thermus sp. FD3041
39	41.2	19.3	1620	13	AAQ22482	Sphingomonas capsu
40	41.2	19.3	2668	13	AAQ22485	S. capsulata iFO12
41	41	19.2	1506	20	AAH76394	Nucleotide sequenc
42	41	19.2	1926	20	AAH82522	Human prostate exp
43	41	19.2	3000	20	AAH84399	
44	41	19.2	18560	21	AAH58472	
45	40.8	19.1	2076	23	ABV22537	

ALIGNMENTS

RESULT 1  
AAS05207  
ID AAS05207 standard; DNA; 214 BP.

AC AAS05207;

DT 07-SEP-2001 (first entry)

DE Mycobacterium chelonae rpoB gene fragment.

KW Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RELP;  
KW PCR-restriction fragment length polymorphism analysis; ds.

OS Mycobacterium chelonae.

PN WO200131061-A1.

PD 03-MAY-2001.

PF 27-OCT-2000; 2000WO-KR01223.

PR 27-OCT-1999; 99KR-0046795.

PA (ERUM-) ERUME BIOTECH CO LTD.

PI Lee H, Park YK, Bai G, Kim S, Kim Y, Cho S, Kim Y, Park HJ;

DR WPI; 2001-300520/31.

PT New DNA fragments from the rpoB gene of mycobacteria, useful for  
PT diagnosis and identification of many mycobacterial species by  
restriction fragment length polymorphism



```

XX PS Claim 1; Page 42; 50pp; English.
XX CC
XX CC The present sequence for Mycobacterium chelonae rpoB gene
XX CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
XX CC various Mycobacterial species. These rpoB gene fragments can be used
XX CC in the diagnosis and identification of Mycobacterium species using a
XX CC novel PCR-restriction fragment length polymorphism analysis (PRA)
XX CC method. The method comprises obtaining a restriction fragment length
XX CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
XX CC amplifying and digesting the DNA fragment from the microorganism to
XX CC be identified and comparing the RFLP patterns from the known rpoB gene
XX CC fragments with the unidentified fragment. The rpoB gene fragments
XX CC are useful to identify a wide range of Mycobacterium species, e.g. for
XX CC diagnosis or to obtain epidemiological and pathogenesis information for
XX CC selection of appropriate therapies, including M. tuberculosis, M. leprae
XX CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected
XX CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene
XX CC fragments is rapid, precise, simple and cost effective (only 1 PCR
XX CC required), and can differentiate between many species in a single
XX CC experiment, including those difficult to distinguish by usual biochemical
XX CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
XX CC detecting specific Mycobacterial species.
XX CC
XX CC Sequence 214 BP; 41 A; 72 C; 72 G; 29 T; 0 other;
XX CC
XX CC Query Match 100.0%; Score 214; DB 22; Length 214;
XX CC Best Local Similarity 100.0%; Pred. No. 9e-39;
XX CC Matches 214; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX CC
XX CC QY 1 TCAAGGAGAAGCGGTACGACCTGGCCCGCGTGGCGGTACAAAGTGAACAAGAGCTGG 60
XX CC Db 1 TCAAGGAGAAGCGGTACGACCTGGCCCGCGTGGCGGTACAAAGTGAACAAGAGCTGG 60
XX CC
XX CC QY 61 GTCTTGGCGGTGCCAACCCGGCTCTGGTACTGCCACACGCTCACCAGGAGACGCTCG 120
XX CC Db 61 GTCTTGGCGGTGCCAACCCGGCTCTGGTACTGCCACACGCTCACCAGGAGACGCTCG 120
XX CC
XX CC QY 121 TCGCCACCATCGGTACCTGGTGGCGCTTGCACGAGGGCCAGACCATGACCGCCCCCG 180
XX CC Db 121 TCGCCACCATCGGTACCTGGTGGCGCTTGCACGAGGGCCAGACCATGACCGCCCCCG 180
XX CC
XX CC QY 181 GCGGCTCGAGTCCCGGTTCAGGTTCAGACGACAT 214
XX CC Db 181 GCGGCTCGAGTCCCGGTTCAGGTTCAGACGACAT 214
XX CC
XX CC RESULT 2
XX CC AAS05221
XX CC ID AAS05221 standard; DNA; 214 BP.
XX CC AC AAS05221;
XX CC XX
XX CC 07-SEP-2001 (first entry)
XX CC
XX CC Mycobacterium abscessus rpoB gene fragment.
XX CC
XX CC Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;
XX CC PCR-restriction fragment length polymorphism analysis; ds.
XX CC
XX CC Mycobacterium abscessus.
XX CC
XX CC WO200131061-A1.
XX CC
XX CC 03-MAY-2001.
XX CC
XX CC 27-OCT-2000; 2000WO-KR01223.
XX CC
XX CC 27-OCT-1999; 99KR-0046795.
XX CC
XX CC (ERUM-) ERUME BIOTECH CO LTD.
XX CC
XX CC Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;

```

```

XX DR WPI; 2001-300520/31.
XX PT
XX PT New DNA fragments from the rpoB gene of mycobacteria, useful for
XX PT diagnosis and identification of many mycobacterial species by
XX PT restriction fragment length polymorphism
XX PS Claim 1; Page 46; 50pp; English.
XX CC
XX CC The present sequence for Mycobacterium abscessus rpoB gene
XX CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
XX CC various Mycobacterial species. These rpoB gene fragments can be used
XX CC in the diagnosis and identification of Mycobacterium species using a
XX CC novel PCR-restriction fragment length polymorphism analysis (PRA)
XX CC method. The method comprises obtaining a restriction fragment length
XX CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
XX CC amplifying and digesting the DNA fragment from the microorganism to
XX CC be identified and comparing the RFLP patterns from the known rpoB gene
XX CC fragments with the unidentified fragment. The rpoB gene fragments
XX CC are useful to identify a wide range of Mycobacterium species, e.g. for
XX CC diagnosis or to obtain epidemiological and pathogenesis information for
XX CC selection of appropriate therapies, including M. tuberculosis, M. leprae
XX CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected
XX CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene
XX CC fragments is rapid, precise, simple and cost effective (only 1 PCR
XX CC required), and can differentiate between many species in a single
XX CC experiment, including those difficult to distinguish by usual biochemical
XX CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
XX CC detecting specific Mycobacterial species.
XX CC
XX CC Sequence 214 BP; 45 A; 69 C; 72 G; 28 T; 0 other;
XX CC
XX CC Query Match 87.3%; Score 186.8; DB 22; Length 214;
XX CC Best Local Similarity 92.1%; Pred. No. 9.6e-33;
XX CC Matches 197; Conservative 0; Mismatches 17; Indels 0; Gaps 0;
XX CC
XX CC QY 1 TCAAGGAGAAGCGGTACGACCTGGCCCGCGTGGCGGTACAAAGTGAACAAGAGCTGG 60
XX CC Db 1 TCAAGGAGAAGCGGTACGACCTGGCCCGCGTGGCGGTACAAAGTGAACAAGAGCTGG 60
XX CC
XX CC QY 61 GTCTTGGCGGTGCCAACCCGGCTCTGGTACTGCCACACGCTCACCAGGAGACGCTCG 120
XX CC Db 61 GCTTGGCGGTGCCAACCCGGCTCTGGTACTGCCACACGCTCACCAGGAGACGCTCG 120
XX CC
XX CC QY 121 TCGCCACCATCGGTACCTGGTGGCGCTTGCACGAGGGCCAGACCATGACCGCCCCCG 180
XX CC Db 121 TCGCCACCATCGGTACCTGGTGGCGCTTGCACGAGGGCCAGACCATGACCGCCCCCG 180
XX CC
XX CC QY 181 GCGGCTCGAGTCCCGGTTCAGGTTCAGACGACAT 214
XX CC Db 181 GCGGCTCGAGTCCCGGTTCAGGTTCAGACGACAT 214
XX CC
XX CC RESULT 3
XX CC AAS05219
XX CC ID AAS05219 standard; DNA; 208 BP.
XX CC AC AAS05219;
XX CC XX
XX CC 07-SEP-2001 (first entry)
XX CC
XX CC Mycobacterium fortuitum rpoB gene fragment.
XX CC
XX CC Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;
XX CC PCR-restriction fragment length polymorphism analysis; ds.
XX CC
XX CC Mycobacterium fortuitum.
XX CC
XX CC WO200131061-A1.
XX CC
XX CC 03-MAY-2001.
XX CC
XX CC 27-OCT-2000; 2000WO-KR01223.

```



DE XX Mycobacterium avium rpoB gene fragment.  
 KW XX Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;  
 KW XX PCR-restriction fragment length polymorphism analysis; ds.  
 XX XX Mycobacterium avium.  
 OS XX WO200131061-A1.  
 PN XX  
 XX 03-MAY-2001.  
 PD XX  
 XX 27-OCT-2000; 2000WO-KR01223.  
 XX 27-OCT-1999; 99KR-0046795.  
 XX (ERUM-) ERUME BIOTECH CO LTD.  
 PA XX  
 XX Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;  
 PT WPI; 2001-300520/31.  
 DR XX  
 XX New DNA fragments from the rpoB gene of mycobacteria, useful for  
 PT diagnosis and identification of many mycobacterial species by  
 PT restriction fragment length polymorphism -  
 XX Claim 1; Page 44; 50pp; English.  
 PS XX  
 XX The present sequence for Mycobacterium avium rpoB gene  
 CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from  
 CC various Mycobacterium species. These rpoB gene fragments can be used  
 CC in the diagnosis and identification of Mycobacterium species using a  
 CC novel PCR-restriction fragment length polymorphism analysis (PRA)  
 CC method. The method comprises obtaining a restriction fragment length  
 CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,  
 CC amplifying and digesting the DNA fragment from the microorganism to  
 CC be identified and comparing the RFLP patterns from the known rpoB gene  
 CC fragments with the unidentified fragment. The rpoB gene fragments  
 CC are useful to identify a wide range of Mycobacterium species, e.g. for  
 CC diagnosis or to obtain epidemiological and pathogenesis information for  
 CC selection of appropriate therapies, including M. tuberculosis, M. leprae  
 CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected  
 CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene  
 CC fragments is rapid, precise, simple and cost effective (only 1 PCR  
 CC required), and can differentiate between many species in a single  
 CC experiment, including those difficult to distinguish by usual biochemical  
 CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for  
 CC detecting specific Mycobacterium species.  
 XX Sequence 208 BP; 44 A; 69 C; 69 G; 26 T; 0 other;  
 SQ  
 Query Match 66.5%; Score 142.4; DB 22; Length 208;  
 Best Local Similarity 82.7%; Pred. No. 6.6e-23;  
 Matches 177; Conservative 0; Mismatches 31; Indels 6; Gaps 1;  
 QY 1 TCAAGGAGAGCGCTACGACCTGGCCGGTGGCCGGTACAGGTGAACAAGAGCTGG 60  
 DB 1 TCAAGGAGAGCGCTACGACCTGGCCGGTGGCCGGTACAGGTGAACAAGAGCTGG 60  
 QY 61 GTCTTGGCGGTGCGCAACCCCGCTCTGTGACTGCCACACAGCTCACCGAGGAGAGCTCG 120  
 DB 61 GCATCGCGCGGTGAGCC-----GATCACAGCTCGAGCGTGCACGAGGAGAGCTCG 114  
 QY 121 TCGCCACCATCGGTACCTGGTGGCGCTGCACGAGGGCGCAGACGATGACCGCCCCCG 180  
 DB 115 TCGCCACCATCGAGTACCTGGTGGCGCTGCACGAGGGTGCAGCCACGATGACCGTCCCGG 174  
 QY 181 CGGCGCTCGAGGTCCCGGTGGAGTGCAGGACAT 214  
 DB 175 CGGCGATCGAGGTCCCGGTGGAGACCGAGACAT 208  
 RESULT 6  
 AAS05218

ID XX AAS05218 standard; DNA; 211 BP.  
 AC XX AAS05218;  
 DT XX 07-SEP-2001 (first entry)  
 XX XX Mycobacterium flavescens rpoB gene fragment.  
 DE XX Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;  
 KW XX PCR-restriction fragment length polymorphism analysis; ds.  
 XX XX Mycobacterium flavescens.  
 OS XX WO200131061-A1.  
 PN XX  
 XX 03-MAY-2001.  
 PD XX  
 XX 27-OCT-2000; 2000WO-KR01223.  
 XX 27-OCT-1999; 99KR-0046795.  
 XX (ERUM-) ERUME BIOTECH CO LTD.  
 PA XX  
 XX Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;  
 PT WPI; 2001-300520/31.  
 DR XX  
 XX New DNA fragments from the rpoB gene of mycobacteria, useful for  
 PT diagnosis and identification of many mycobacterial species by  
 PT restriction fragment length polymorphism -  
 XX Claim 1; Page 45; 50pp; English.  
 PS XX  
 XX The present sequence for Mycobacterium flavescens rpoB gene  
 CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from  
 CC various Mycobacterium species. These rpoB gene fragments can be used  
 CC in the diagnosis and identification of Mycobacterium species using a  
 CC novel PCR-restriction fragment length polymorphism analysis (PRA)  
 CC method. The method comprises obtaining a restriction fragment length  
 CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,  
 CC amplifying and digesting the DNA fragment from the microorganism to  
 CC be identified and comparing the RFLP patterns from the known rpoB gene  
 CC fragments with the unidentified fragment. The rpoB gene fragments  
 CC are useful to identify a wide range of Mycobacterium species, e.g. for  
 CC diagnosis or to obtain epidemiological and pathogenesis information for  
 CC selection of appropriate therapies, including M. tuberculosis, M. leprae  
 CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected  
 CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene  
 CC fragments is rapid, precise, simple and cost effective (only 1 PCR  
 CC required), and can differentiate between many species in a single  
 CC experiment, including those difficult to distinguish by usual biochemical  
 CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for  
 CC detecting specific Mycobacterium species.  
 XX Sequence 211 BP; 49 A; 65 C; 70 G; 27 T; 0 other;  
 SQ  
 Query Match 65.6%; Score 140.4; DB 22; Length 211;  
 Best Local Similarity 81.8%; Pred. No. 1.8e-22;  
 Matches 175; Conservative 0; Mismatches 36; Indels 3; Gaps 1;  
 QY 1 TCAAGGAGAGCGCTACGACCTGGCCGGTGGCCGGTACAGGTGAACAAGAGCTGG 60  
 DB 1 TCAAGGAGAGCGCTACGACCTGGCCGGTGGCCGGTACAGGTGAACAAGAGCTGG 60  
 QY 61 GTCTTGGCGGTGCCAACCCCGCTCTGTGACTGCCACACAGCTCACCGAGGAGAGCTCG 120  
 DB 61 GCATCGCGCGGTGAGCC-----GATCACAGCTCGAGCGTGCACGAGGAGAGCTCG 117  
 QY 121 TCGCCACCATCGGTACCTGGTGGCGCTGCACGAGGGCGCAGACGATGACCGCCCCCG 180  
 DB 118 TCGCCACCATCGAGTACCTGGTGGCGCTGCATCAGGCGGACAAAGACGATGACCGTCCCGG 177  
 QY 181 CGGCGCTCGAGGTCCCGGTGGAGTGCAGGACAT 214

Db	178	GTGAGTCGAGTGCCTCGAGTCGACGACAT	211
RESULT 7			
AAAS05201			
ID	AAAS05201	standard; DNA; 208 BP.	
XX	AC	AAAS05201;	
XX	XX	07-SEP-2001 (first entry)	
DT	DT		
XX	XX	Mycobacterium gordonae type I rpoB gene fragment.	
DE	DE		
XX	XX	Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;	
KW	KW	PCR-restriction fragment length polymorphism analysis; ds.	
XX	XX		
OS	OS	Mycobacterium gordonae type I.	
XX	XX	WO200131061-A1.	
PN	PN		
XX	XX	03-MAY-2001.	
PD	PD		
XX	XX		
PF	PF	27-OCT-2000; 2000WO-KR01223.	
XX	XX		
PR	PR	27-OCT-1999; 99KR-0046795.	
XX	XX	(ERUN-) ERUME BIOTECH CO LTD.	
PA	PA		
XX	XX		
PI	PI	Lee H, Park YK, Bal G, Kim S, Cho S, Kim Y, Park HJ;	
XX	XX	WPI; 2001-300520/31.	
DR	DR		
XX	XX		
PT	PT	New DNA fragments from the rpoB gene of mycobacteria, useful for	
PT	PT	diagnosis and identification of many mycobacterial species by	
PT	PT	restriction fragment length polymorphism -	
XX	XX		
PS	PS	Claim 1: Page 40: 50pp; English.	
CC	CC		
XX	XX	The present sequence for Mycobacterium gordonae type I rpoB gene	
CC	CC	fragment is 1 of 24 rpoB gene fragments (AAAS05201-AAAS05224) from	
CC	CC	various Mycobacterium species. These rpoB gene fragments can be used	
CC	CC	in the diagnosis and identification of Mycobacterium species using a	
CC	CC	novel PCR-restriction fragment length polymorphism analysis (PRA)	
CC	CC	method. The method comprises obtaining a restriction fragment length	
CC	CC	polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,	
CC	CC	amplifying and digesting the DNA fragment from the microorganism to	
CC	CC	be identified and comparing the RFLP patterns from the known rpoB gene	
CC	CC	fragments with the unidentified fragment. The rpoB gene fragments	
CC	CC	are useful to identify a wide range of Mycobacterium species, e.g. for	
CC	CC	diagnosis or to obtain epidemiological and pathogenesis information for	
CC	CC	selection of appropriate therapies, including M. tuberculosis, M. leprae	
CC	CC	and non-tuberculous mycobacteria (NTM) encountered in subjects infected	
CC	CC	with human immunodeficiency virus (HIV). Analysis of the rpoB gene	
CC	CC	fragments is rapid, precise, simple and cost effective (only 1 PCR	
CC	CC	required), and can differentiate between many species in a single	
CC	CC	experiment, including those difficult to distinguish by usual biochemical	
CC	CC	tests. Also described are oligonucleotide probes (AAAS05227-AAAS05242) for	
CC	CC	detecting specific Mycobacterium species.	
XX	XX		
SQ	SQ	Sequence 208 BP; 46 A; 71 C; 65 G; 26 T; 0 other;	
Query Match	65.0%	Score 139.2; DB 22; Length 208;	
Best Local Similarity	81.8%	Pred. No. 3.4e-22;	
Matches 175; Conservative	0; Mismatches 33; Indels 6; Gaps 1;		
00y	1	TCAAGGAGAAGCGGTACGACCTGGCCCGCGTGGGCGCGGTACAAGGTGAACAAGAAGCTGG	60
Db	1	TCAAGGAGAAGCGGTACGACCTGGCCCGCGGTAGGCGCGCTACAAGGTCACAGAAGCTCG	60
00y	61	GTCCTGGCGGTGCCAACCCGGCTCTGCTGACTGCCACCAGCGCTACCCGAGGAAGAGCTCG	120
Db	61	GCTTCGACGTCCGCG-----GATCCGATCACAGCTCCAGCTTCAGCTCCGAGGAGAGCTCG	114

QY	1	TCAAGGAGACGGGTACGACTGCCTGCCCGGTGGCCGGTACAAGGTGAACAAGAAGCTGG	60
Db	1	TCAAGGAGAACGGTACGAGCTGGCTCGGTGGTTCGATCAGGTTGAAAGGAGGCTGG	60
QY	61	GTCATTGGCGGTGCCAACC CGGCTCTGGTGACTGCCACCAACGCTCACCGAGGAGAGCTGG	120
Db	61	GCCTGAAGCGGCCAGCC-----CATCACCAGGTCGACGCTACCGGAGGAGAGCTGG	114
QY	121	TGCCACCATTCCGGTAAGCTGGTCGGCTGCACGAGGGCCAGACCAAGATGACCGCCCCCG	180
Db	115	TGCGCCACCATCGAATACTTGGTCCGCTTCGACGAGGGCCAGACCGGATGACCGCTCCGG	174
QY	181	GGGGCTTCAGGTCGCGGTTCGAGGTTCGACGACAT	214
Db	175	GGGGTGTTCGAGGTGCGCGTTCGAGACGACGACAT	208

The present sequence for *Mycobacterium gordonae* type II rpoB gene fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from various *Mycobacterial* species. These rpoB gene fragments can be used in the diagnosis and identification of *Mycobacterium* species using a novel PCR-restriction fragment length polymorphism analysis (PRA) method. The method comprises obtaining a restriction fragment length polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating, amplifying and digesting the DNA fragment from the microorganism to be identified and comparing the RFLP patterns from the known rpoB gene fragments with the unidentified fragment. The rpoB gene fragments are useful to identify a wide range of *Mycobacterium* species, e.g. for diagnosis or to obtain epidemiological and pathogenesis information for selection of appropriate therapies, including *M. tuberculosis*, *M. leprae* and non-tuberculous mycobacteria (NTM) encountered in subjects infected with human immunodeficiency virus (HIV). Analysis of the rpoB gene fragments is rapid, precise, simple and cost effective (only 1 PCR

The present sequence for *Mycobacterium ulcerans* rpoB gene fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from various *Mycobacterial* species. These rpoB gene fragments can be used in the diagnosis and identification of *Mycobacterium* species using a novel PCR-restriction fragment length polymorphism analysis (PRA) method. The method comprises obtaining a restriction fragment length polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating, amplifying and digesting the DNA fragment from the microorganism to be identified and comparing the RFLP patterns from the known rpoB gene fragments with the unidentified fragment. The rpoB gene fragments are useful to identify a wide range of *Mycobacterium* species, e.g. for diagnosis or to obtain epidemiological and pathogenesis information for selection of appropriate therapies, including *M. tuberculosis*, *M. leprae* and non-tuberculous *Mycobacteria* (NTM) encountered in subjects infected with human immunodeficiency virus (HIV). Analysis of the rpoB gene fragments is rapid, precise, simple and cost effective (only 1 PCR required), and can differentiate between many species in a single experiment, including those difficult to distinguish by usual biochemical tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for detecting specific *Mycobacterial* species.

Sequence 208 BP; 45 A; 68 C; 67 G; 28 T; 0 other;

CC required), and can differentiate between many species in a single  
CC experiment, including those difficult to distinguish by usual biochemical  
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for  
CC detecting specific Mycobacterium species.

XX  
SQ Sequence 208 BP; 45 A; 62 C; 73 G; 28 T; 0 other;

Query Match 63.6%; Score 136; DB 22; Length 208;  
Best Local Similarity 80.8%; Pred. No. 1.7e-21;  
Matches 173; Conservative 0; Mismatches 35; Indels 6; Gaps 1;

QY 1 TCAAGGAGAAGCGCTACGACCTGGCGCGTGGCGGTGACAAAGTGAAAGAGCTGG 60  
DB 1 TCAAGGAGAAGCGCTACGACCTGGCGCGGTGGCGGTGACAAAGTGAAAGAGCTGG 60  
QY 61 GTCTTGGCGGTGCAACCGCGCTCTGGTACCTGACACACCGCTACCGAGGAAGCGTCG 120  
DB 61 GTCTTGGCGGTGCAACCGCGCTCTGGTACCTGACACACCGCTACCGAGGAAGCGTCG 114  
QY 121 TCGCCACCATCGGCTACCTGGTGGCGCTGCACGAGGGCGACACACGATGACCGCCCG 180  
DB 115 TAGCCACCATCGAGTACCTGGTGGCGCTGCACGAGGGTCAGTCGCGGATGACGGTTCCG 174  
QY 181 GCGGCTCGAGTCCGCGTGGAGTCGACGACAT 214  
DB 175 GCGGCGCGAGGTGGCGGTGGAGACCGACGACAT 208

## RESULT 11

AAS05217  
ID AAS05217 standard; DNA; 208 BP.

XX  
AC AAS05217;

XX  
DT 07-SEP-2001 (first entry)

XX  
DE Mycobacterium celatum rpoB gene fragment.

XX  
KW Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;  
PCR-restriction fragment length polymorphism analysis; ds.

XX  
OS Mycobacterium celatum.

XX  
PN WO200131061-A1.

XX  
PD 03-MAY-2001.

XX  
PF 27-OCT-2000; 2000WO-KR01223.

XX  
PR 27-OCT-1999; 99KR-0046795.

XX  
PA (ERUM-) ERUME BIOTECH CO LTD.

XX  
PI Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;

XX  
DR WPI; 2001-300520/31.

XX  
PT New DNA fragments from the rpoB gene of mycobacteria, useful for  
diagnosis and identification of many mycobacterial species by  
restriction fragment length polymorphism.

XX  
PS Claim 1; Page 45; 50pp; English.

XX  
CC The present sequence for Mycobacterium celatum rpoB gene  
fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from  
various Mycobacterium species. These rpoB gene fragments can be used  
in the diagnosis and identification of Mycobacterium species using a  
novel PCR-restriction fragment length polymorphism analysis (PRA)  
method. The method comprises obtaining a restriction fragment length  
polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,  
amplifying and digesting the DNA fragment from the microorganism to  
be identified and comparing the RFLP patterns from the known rpoB gene  
fragments with the unidentified fragment. The rpoB gene fragments

CC are useful to identify a wide range of Mycobacterium species, e.g. for  
CC diagnosis or to obtain epidemiological and pathogenesis information for  
CC selection of appropriate therapies, including M. tuberculosis, M. leprae  
CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected  
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene  
CC fragments is rapid, precise, simple and cost effective (only 1 PCR  
CC required), and can differentiate between many species in a single  
CC experiment, including those difficult to distinguish by usual biochemical  
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for  
CC detecting specific Mycobacterium species.

XX  
SQ Sequence 208 BP; 48 A; 71 C; 64 G; 25 T; 0 other;

Query Match 62.8%; Score 134.4; DB 22; Length 208;  
Best Local Similarity 80.4%; Pred. No. 3.9e-21;  
Matches 172; Conservative 0; Mismatches 36; Indels 6; Gaps 1;

QY 1 TCAAGGAGAAGCGCTACGACCTGGCGCGGTGGCGGTGACAAAGTGAAAGAGCTGG 60  
DB 1 TCAAGGAGAAGCGCTACGACCTGGCGCGGTGGCGGTGACAAAGTGAAAGAGCTGG 60  
QY 61 GTCTTGGCGGTGCCAACCGCGCTCTGGTACCTGACACACCGCTACCGAGGAAGCGTCG 120  
DB 61 GCCT-----GAACACCGCGTCCCGATCACGACGACCTCTGACCGAGAGCGTCG 114  
QY 121 TCGCCACCATCGGCTACCTGGTGGCGCTGCACGAGGGCGACACGATGACCGCCCG 180  
DB 115 TCGCCACCATCGAGTACCTGGTGGCGCTGCACGAGGGCGACACGATGACCGCCCG 174  
QY 181 GCGGCTCGAGTCCGCGTGGAGTCGACGACAT 214  
DB 175 GCGGAGTCGAGTGGCGGTGGAAACCGACGACAT 208

## RESULT 12

AAS05203  
ID AAS05203 standard; DNA; 208 BP.

XX  
AC AAS05203;

XX  
DT 07-SEP-2001 (first entry)

XX  
DE Mycobacterium gordonae type III rpoB gene fragment.

XX  
KW Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;  
PCR-restriction fragment length polymorphism analysis; ds.

XX  
OS Mycobacterium gordonae type III.

XX  
PN WO200131061-A1.

XX  
PD 03-MAY-2001.

XX  
PF 27-OCT-2000; 2000WO-KR01223.

XX  
PR 27-OCT-1999; 99KR-0046795.

XX  
PA (ERUM-) ERUME BIOTECH CO LTD.

XX  
PI Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;

XX  
DR WPI; 2001-300520/31.

XX  
PT New DNA fragments from the rpoB gene of mycobacteria, useful for  
diagnosis and identification of many mycobacterial species by  
restriction fragment length polymorphism.

XX  
PS Claim 1; Page 41; 50pp; English.

XX  
CC The present sequence for Mycobacterium gordonae type III rpoB gene  
fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from  
various Mycobacterium species. These rpoB gene fragments can be used  
in the diagnosis and identification of Mycobacterium species using a

CC novel PCR-restriction fragment length polymorphism analysis (PRA)  
 CC method. The method comprises obtaining a restriction fragment length  
 CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,  
 CC amplifying and digesting the DNA fragment from the microorganism to  
 CC be identified and comparing the RFLP patterns from the known rpoB gene  
 CC fragments with the unidentified fragment. The rpoB gene fragments  
 CC are useful to identify a wide range of Mycobacterium species, e.g. for  
 CC diagnosis or to obtain epidemiological and pathogenesis information for  
 CC selection of appropriate therapies, including M. tuberculosis, M. leprae  
 CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected  
 CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene  
 CC fragments is rapid, precise, simple and cost effective (only 1 PCR  
 CC required), and can differentiate between many species in a single  
 CC experiment, including those difficult to distinguish by usual biochemical  
 CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for  
 CC detecting specific Mycobacterial species.

XX Sequence 208 BP; 46 A; 69 C; 63 G; 30 T; 0 other;

Query Match 62.1%; Score 132.8; DB 22; Length 208;  
 Best Local Similarity 79.9%; Pred. No. 8.9e-21;  
 Matches 171; Conservative 0; Mismatches 37; Indels 6; Gaps 1;

QY 1 TCAAGGAGAGCGGTACGACCTGGCCCGGTGGCGGTGCTGACTCCACCGCTACAGGTCAACAGAGCTGG 60  
 Db 1 TCAAGGAGAGCGGTACGACCTGGCCCGGTGGCGGTGCTGACTCCACCGCTACAGGTCAACAGAGCTGG 60  
 QY 61 GTCTTGGCGGTGCAACCGCGCTGTGTGACTCCACCGCTACAGGTCAACAGAGCTGG 120  
 Db 61 GCGTCCAGCTCGGC-----GATCCGATCACCGCTCCAGCTGACCGGAGAGAGCTGG 114  
 QY 121 TCGCCACCATCGGTACCTGGTGGCTGCGCAGGAGGCGCCAGACCGATGACCGCCCGG 180  
 Db 115 TCGCCACCATCGGTACCTGGTGGCTGCGCAGGAGGCGCCAGACCGATGACCGCCCGG 174  
 QY 181 GCGGCTCGAGGTCCCGGTGCGAGGTGCGACGACAT 214  
 Db 175 GCGGACCGAGGTTCGCGTGGAGACCGACGACAT 208

RESULT 13  
 AAS05208

ID AAS05208 standard; DNA; 208 BP.

AC AAS05208;

DT 07-SEP-2001 (first entry)

DE Mycobacterium kansasii rpoB gene fragment.

KW Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;  
 KW PCR-restriction fragment length polymorphism analysis; ds.

OS Mycobacterium kansasii.

PN WO200131061-A1.

PD 03-MAY-2001.

PF 27-OCT-2000; 2000WO-KR01223.

PR 27-OCT-1999; 99KR-0046795.

PA (ERUM-) ERUME BIOTECH CO LTD.

XX Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;  
 PI WPI: 2001-300520/31.

XX New DNA fragments from the rpoB gene of mycobacteria, useful for  
 PT diagnosis and identification of many mycobacterial species by  
 PT restriction fragment length polymorphism

XX

PS

XX Claim 1; Page 42; 50pp; English.

XX

CC The present sequence for Mycobacterium kansasii rpoB gene  
 CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from  
 CC various Mycobacterial species. These rpoB gene fragments can be used  
 CC in the diagnosis and identification of Mycobacterium species using a  
 CC novel PCR-restriction fragment length polymorphism analysis (PRA)  
 CC method. The method comprises obtaining a restriction fragment length  
 CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,  
 CC amplifying and digesting the DNA fragment from the microorganism to  
 CC be identified and comparing the RFLP patterns from the known rpoB gene  
 CC fragments with the unidentified fragment. The rpoB gene fragments  
 CC are useful to identify a wide range of Mycobacterium species, e.g. for  
 CC diagnosis or to obtain epidemiological and pathogenesis information for  
 CC selection of appropriate therapies, including M. tuberculosis, M. leprae  
 CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected  
 CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene  
 CC fragments is rapid, precise, simple and cost effective (only 1 PCR  
 CC required), and can differentiate between many species in a single  
 CC experiment, including those difficult to distinguish by usual biochemical  
 CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for  
 CC detecting specific Mycobacterial species.

XX Sequence 208 BP; 51 A; 65 C; 65 G; 27 T; 0 other;

Query Match 62.1%; Score 132.8; DB 22; Length 208;  
 Best Local Similarity 79.9%; Pred. No. 8.9e-21;  
 Matches 171; Conservative 0; Mismatches 37; Indels 6; Gaps 1;

QY 1 TCAAGGAGAGCGGTACGACCTGGCCCGGTGGCGGTGCTGACTCCACCGCTACAGGTCAACAGAGCTGG 60  
 Db 1 TCAAGGAGAGCGGTACGACCTGGCCCGGTGGCGGTGCTGACTCCACCGCTACAGGTCAACAGAGCTGG 60  
 QY 61 GTCTTGGCGGTGCAACCGCGCTGTGTGACTCCACCGCTACAGGTCAACAGAGCTGG 120  
 Db 61 GCCTCGAGCTCGGC-----GAACACCAATCATCCGATCACCGGAGAGAGCTGG 114  
 QY 121 TCGCCACCATCGGTACCTGGTGGCTGCGCAGGAGGCGCCAGACCGATGACCGCCCGG 180  
 Db 115 TCGCCACCATCGGTACCTGGTGGCTGCGCAGGAGGCGCCAGACCGATGACCGCCCGG 174  
 QY 181 GCGGCTCGAGGTCCCGGTGCGAGGTGCGACGACAT 214  
 Db 175 GCGGGTCTGAGGTGCGCGTGGAGAACCGACGACAT 208

RESULT 14  
 AAS05224

ID AAS05224 standard; DNA; 208 BP.

AC AAS05224;

DT 07-SEP-2001 (first entry)

DE Mycobacterium xenopi rpoB gene fragment.

KW Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;  
 KW PCR-restriction fragment length polymorphism analysis; ds.

OS Mycobacterium xenopi.

PN WO200131061-A1.

PD 03-MAY-2001.

PF 27-OCT-2000; 2000WO-KR01223.

PR 27-OCT-1999; 99KR-0046795.

PA (ERUM-) ERUME BIOTECH CO LTD.

XX Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;







GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 12, 2002, 16:50:46 : Search time 654.37 Seconds  
(without alignments)  
9517.553 Million cell updates/sec

Title: US-09-697-123b-7

Perfect score: 214

Sequence: 1 tcaaggagaagcgctacgac.....ccggtcgaggtcgacgacat 214

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl.\*

1: gb\_ba.\*

2: gb\_htg.\*

3: gb\_in.\*

4: gb\_om.\*

5: gb\_ov.\*

6: gb\_pat.\*

7: gb\_ph.\*

8: gb\_pl.\*

9: gb\_pr.\*

10: gb\_ro.\*

11: gb\_sts.\*

12: gb\_sy.\*

13: gb\_un.\*

14: gb\_vi.\*

15: em\_ba.\*

16: em\_fun.\*

17: em\_hum.\*

18: em\_in.\*

19: em\_mu.\*

20: em\_om.\*

21: em\_or.\*

22: em\_ov.\*

23: em\_pat.\*

24: em\_ph.\*

25: em\_pl.\*

26: em\_ro.\*

27: em\_sts.\*

28: em\_un.\*

29: em\_vi.\*

30: em\_htg\_hum.\*

31: em\_htg\_inv.\*

32: em\_htg\_other.\*

33: em\_htg\_mus.\*

34: em\_htg\_pln.\*

35: em\_htg\_rtd.\*

36: em\_htg\_mam.\*

37: em\_htg\_vrt.\*

38: em\_sy.\*

39: em\_htgo\_hum.\*

40: em\_htgo\_mus.\*

41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	132.8	62.1	3752	1	MSU24494	U24494 Mycobacteri
2	124.8	58.3	610	1	MTU318818	AJ318818 Mycobacte
3	124.8	58.3	610	1	MTU318819	AJ318819 Mycobacte
4	124.8	58.3	616	1	MTU318813	AJ318813 Mycobacte
5	124.8	58.3	618	1	MTU318815	AJ318815 Mycobacte
6	124.8	58.3	618	1	MTU318817	AJ318817 Mycobacte
7	124.8	58.3	633	1	MTU318814	AJ318814 Mycobacte
8	124.8	58.3	637	1	MTU318816	AJ318816 Mycobacte
9	124.8	58.3	639	1	MTU318821	AJ318821 Mycobacte
10	124.8	58.3	970	6	150706	150706 Sequence 1
11	124.8	58.3	3534	6	AX111339	AX111339 Sequence
12	124.8	58.3	3853	1	MTU12205	MTU12205 Mycobacteri
13	124.8	58.3	5084	1	MSGRPO8	L27989 Mycobacteri
14	124.8	58.3	19352	1	AE006964	AE006964 Mycobacte
15	124.8	58.3	19770	1	MTU31376	295972 Mycobacteri
16	104	48.6	3447	6	AR067447	AR067447 Sequence
17	104	48.6	37617	1	MLB1790G	Z14314 M.leprae ge
18	104	48.6	348950	1	MLEPRTN7	AL583923 Mycobacte
19	103.2	48.2	3941	1	AF242549	AF242549 Mycolato
20	88.4	41.3	3495	6	AX120631	AX120631 Sequence
21	88.4	41.3	328050	1	AP005375	AP005375 Corynebac
22	88.4	41.3	34980	6	AX127144	AX127144 Sequence
23	84	39.3	32923	1	SCD82	AL160431 Streptomy
24	47	22.0	11103	1	AE001944	AE001944 Deinococc
25	46.8	21.9	7854	1	AX024345	AX024345 Sequence
26	46.8	21.9	7854	6	AX024238	AX024238 Sequence
27	46.8	21.9	47713	1	AX024320	AX024320 Sequence
28	46.8	21.9	47713	6	AX024213	AX024213 Sequence
29	46	21.5	3350	1	SCAPTRELA	X87267 S.coelicolo
30	44.8	20.9	101365	2	AF004748	AF004748 Oryza sat
31	44.8	20.9	152263	2	AF005412	AF005412 Oryza sat
32	44.6	20.8	329100	1	SME591787	AL591787 Sinorhizo
33	44.4	20.7	4068	1	SCPPGPP	X95250 S.coelicolo
34	44.4	20.7	38640	1	SCL2	AL137778 Streptomy
35	44	20.6	30561	1	SC1F2	AL031350 Streptomy
36	44	20.6	34023	1	SC2G2	AL445963 Streptomy
37	43.6	20.4	1810	1	AF222752	AF222752 Bradyrhiz
38	43.6	20.4	12277	1	AF340166	AF340166 Streptomy
39	43.6	20.4	37159	2	AC100259	AC100259 Mus muscu
40	43.6	20.4	134188	2	AC120983	AC120983 Oryza sat
41	43.4	20.3	4746	1	ABHISHAFE	X61207 A.brasilens
42	43.2	20.2	17849	1	AE007096	AE007096 Mycobacte
43	43.2	20.2	33818	1	MTU3159	Z83863 Mycobacteri
44	43	20.1	2904	1	SREFRPS	Y14317 Streptomyce
45	43	20.1	118733	2	AP003885	AP003885 Oryza sat

ALIGNMENTS

RESULT 1  
MSU24494  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

MSU24494  
Mycobacterium smegmatis DNA polymerase (rpoB) gene, complete cds.  
U24494  
U24494.1 GI:790347

Mycobacterium smegmatis.  
Mycobacterium smegmatis.  
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
Actinomycetales; Corynebacterineae; Mycobacteriaceae;  
Mycobacterium.  
1 (bases 1 to 3752)

REFERENCE  
AUTHORS  
TITLE  
Hetherington,S.V., Watson,A.S. and Patrick,C.C.  
Sequence and analysis of the rpoB gene of Mycobacterium smegmatis

JOURNAL Antimicrob. Agents Chemother. 39 (9), 2164-2166 (1995)  
MEDLINE 96050766  
PUBMED 8540740  
REFERENCE 2 (bases 1 to 3752)  
AUTHORS Hetherington,S.V.  
TITLE Direct Submission  
JOURNAL Submitted (11-APR-1995) Seth V. Hetherington, Infectious Diseases,  
St. Jude Children's Research Hospital, 332 N. Lauderdale, Memphis,  
TN 38101, USA

FEATURES  
source Location/Qualifiers  
1..3752  
/organism="Mycobacterium smegmatis"  
/db\_xref="taxon:1772"  
194..3703  
/gene="rpoB"  
/gene="rpoB"  
/codon\_start=1  
/evidence=experimental  
/transl\_table=11  
/product="DNA polymerase"  
/protein\_id="AAA91426.2"  
/db\_xref="GI:7144498"  
LDVQDSFEMLVGSRVROAIDRGEENPVGLAEVIAELSPTEDFSGMSLSFSDPR  
FDEVKASDFECKDKDMTYAAPLFVTAEFINNNGEIKSQTFVEMGDFPMTEKGTFFIN  
GTERVVSOLVRSFVYFDETDKSTKTLHVKVIFPGANLEFVDKRDVTGVRID  
PKRKQPVTLKALGNCTNEQIVRFGESEIMMGTLEKDTSGTDEALDLYRKLRPGE  
PPTKSAQTLLENLFFKRYKRYDLARVYKVKLGNAGKPISTLTEDVYATIE  
YLRLHEGQTSMTVPGGVFAVESHDIDHFNRLRTVGLIQNIRVLSRMERVVR  
ERMVTDQVEAITPOTLINIRPVAAIKKEFTGTSOLFQMDONNPLSLGTHKRLSALG  
PGGLSRAGLEVLRDHPHVRMCPETETPEGNIGLIGLSVYARVNPFGFETPYR  
KVVSRRVTDQDLYLTADEEDPHVVAQANSTEDENGRTEREDVVRKKGVEFVSADQ  
VYMDVSRQHVSVATAMPIFLEHDDANRALMGAMQRAVPLVRSEAPLVGTGMELR  
AAIDAVTSKHTQVIEVSADYITVMADPGTQSYRLKRFARNSNCTCANQRPVDA  
QROAGOVIAQDQCTGEONMAGLNLLVAMPWEGHNYEDAIIISRLVEEDVLTSP  
HIEHEIDARQTLKAGBEITRIVPKLSDLEVALDDEPRIVRGAENVDRGILVKGVTI  
KGTELPPEERLLRAIFGEKAREVRDTSKVPHGSEKVGIRVFESREDDDELPGVGN  
ELVYVVAQRKISDQKLAHNGKVGIGKILPVEDMPGLDGTVPDVIILNTHGCRV  
LVNTOGILETHLGHKAGKAGNIDVLAGVPWRASKLPEELISAPADSTVATPVFDGAQE  
GELAGLLGSTLPNRDGEVYNADKATLFDGRSGSEFPFYPVTVGYMILKHLHVD  
KIHARSTGPSYMTQOPLGGKAQDFGQRFGEKMECMAQYGAAYTLQELLTIKSDDTV  
GRVKVYEAIVKGENIPEPGIPESFKVLLKELQSLCNLNEVLSSDGRAIENRDRGDDDL  
ERAAANGINLSRNESASVEDLA"

BASE COUNT 706 a 1225 c 1210 g 611 t

Query Match 62.1%; Score 132.8; DB 1; Length 3752;  
Best Local Similarity 79.9%; Pred. No. 1.7e-13;  
Matches 171; Conservative 0; Mismatches 37; Indels 6; Gaps 1;

Qy 1 TCAAGGAGAGCGCTACGACCTGGCCGCGGTGGCGGTACAGGTGACCAAGAGCTGG 60  
|||||  
Db 1068 TCAAGGAGAGCGCTACGACCTGGCCGCGGTGGCGGTACAGGTGACCAAGAGCTGG 1127  
Qy 61 GTCTTGGCGGTGCCAACCCGGCTCTGTGACTGCCACACCGCTCACCGAGGAAGAGCTGG 120  
|||  
Db 1128 GCCTGAACGCGGCAAGCC-----GATCACCAGCTCGACCTGACCGAAGAGAGCTGG 1181  
Qy 121 TCGCCACCATCGGTACTGTGGCCCTGACGAGGGCCAGACCATGATACCCGCCCG 180  
|||  
Db 1182 TCGCAGCATCGATACCTGTGGTCTGTCAGAGGGTTCAGACCTCGATGATACCCGCCG 1241  
Qy 181 GCGGCTCGAGTCCCGCTCGAGTTCGAGCTCGACGACAT 214  
|||  
Db 1242 GTGGCGTCTGAGTTCGCCCTCGAGTTCGACGACGACAT 1275

RESULT 2  
LOCUS MTU318818 610 bp DNA circular BCT 09-AUG-2002  
DEFINITION Mycobacterium tuberculosis partial rpoB gene for RNA polymerase  
beta subunit, isolate 1415-97.

ACCESSION AJ318818  
VERSION AJ318818.1 GI:22208412  
KEYWORDS RNA polymerase beta subunit; rpoB gene.  
SOURCE Mycobacterium tuberculosis.  
ORGANISM Mycobacterium tuberculosis.  
REFERENCE 1  
AUTHORS Herrera,L., Jimenez,M.S. and Saez,J.A.  
TITLE Molecular analysis of rifampin-resistant Mycobacterium tuberculosis isolated in Spain (1996-2001). Description of new alleles into rpoB gene and review  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 610)  
AUTHORS Herrera,L.  
TITLE Direct Submission  
JOURNAL Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5, Majadahonda. Madrid. 28220, SPAIN

FEATURES  
source Location/Qualifiers  
1..610  
/organism="Mycobacterium tuberculosis"  
/isolate="1415-97"  
/db\_xref="taxon:1773"  
1..610  
/gene="rpoB"  
/gene="rpoB"  
/codon\_start=1  
/transl\_table=11  
/product="RNA polymerase beta subunit"  
/protein\_id="CAC87035.1"  
/db\_xref="GI:22208413"  
/translation="LDIYKRLPGEPPTKESQTLLENLFFKRYKRYDLARVYKVK  
KLGLHVGEPITSTLTEDVATIEYLRLHEGQTSMTVPGGVFVETDDIDHFGNR  
RLRTVGLLQNIQIRVMSRMERVIRMTQDVEAITPOTLINIRPVVAAIKKEFTGTS  
QFSQMGQNPLSLGTHKRLSALGPGGLSRAGLEVLRDHPHVR"

BASE COUNT 122 a 191 c 202 g 95 t

Query Match 58.3%; Score 124.8; DB 1; Length 610;  
Best Local Similarity 77.6%; Pred. No. 4.9e-12;  
Matches 166; Conservative 0; Mismatches 42; Indels 6; Gaps 1;

Qy 1 TCAAGGAGAGCGCTACGACCTGGCCGCGGTGGCGGTACAGGTGACCAAGAGCTGG 60  
|||||  
Db 80 TCAAGGAGAGCGCTACGACCTGGCCGCGGTGGCGGTACAGGTGACCAAGAGCTGG 139  
Qy 61 GTCTTGGCGGTGCCAACCCGGCTCTGTGACTGCCACACCGCTCACCGAGGAAGAGCTGG 120  
|||  
Db 140 GGCTGCAATGTCGCGGAGCC-----CATCAGCTCGTGCAGCTGACCGAAGAGAGCTGG 193  
Qy 121 TCGCCACCATCGGTACTGTGGCTGTCAGGAGGCCAGACCATGATACCCGCCCG 180  
|||||  
Db 194 TGGCCACCATCGAATATCTGTCGCTTGCACGAGGGTTCAGACCATGATGCTTCCCG 253  
Qy 181 GCGGCTCGAGTCCCGGTGAGGTTCGAGCTGACGACAT 214  
|||||  
Db 254 GCGGCTCGAGGTGCGCGGTGGAACCGACGACAT 287

RESULT 3  
LOCUS MTU318819 610 bp DNA circular BCT 09-AUG-2002  
DEFINITION Mycobacterium tuberculosis partial rpoB gene for RNA polymerase  
beta subunit, isolate 1417-97.

ACCESSION AJ318819  
VERSION AJ318819.1 GI:22208414  
KEYWORDS RNA polymerase beta subunit; rpoB gene.  
SOURCE Mycobacterium tuberculosis.  
ORGANISM Mycobacterium tuberculosis.  
REFERENCE 1  
AUTHORS Herrera,L., Jimenez,M.S. and Saez,J.A.  
TITLE Molecular analysis of rifampin-resistant Mycobacterium tuberculosis isolated in Spain (1996-2001). Description of new alleles into rpoB gene and review  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 610)  
AUTHORS Herrera,L.  
TITLE Direct Submission  
JOURNAL Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5, Majadahonda. Madrid. 28220, SPAIN

FEATURES  
source Location/Qualifiers  
1..610  
/organism="Mycobacterium tuberculosis"  
/isolate="1415-97"  
/db\_xref="taxon:1773"  
1..610  
/gene="rpoB"  
/gene="rpoB"  
/codon\_start=1  
/transl\_table=11  
/product="RNA polymerase beta subunit"  
/protein\_id="CAC87035.1"  
/db\_xref="GI:22208413"  
/translation="LDIYKRLPGEPPTKESQTLLENLFFKRYKRYDLARVYKVK  
KLGLHVGEPITSTLTEDVATIEYLRLHEGQTSMTVPGGVFVETDDIDHFGNR  
RLRTVGLLQNIQIRVMSRMERVIRMTQDVEAITPOTLINIRPVVAAIKKEFTGTS  
QFSQMGQNPLSLGTHKRLSALGPGGLSRAGLEVLRDHPHVR"

BASE COUNT 122 a 191 c 202 g 95 t

Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium tuberculosis complex.

REFERENCE  
AUTHORS  
TITLE  
Herrera, L., Jimenez, M.S. and Saez, J.A.  
Molecular analysis of rifampin-resistant Mycobacterium tuberculosis isolated in Spain (1996-2001). Description of new alleles into rpoB gene and review

JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
Unpublished  
2 (bases 1 to 610)

JOURNAL  
TITLE  
Direct Submission  
Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5, Majadahonda. Madrid. 28220, SPAIN

FEATURES  
source  
1. .610  
Location/Qualifiers

/organism="Mycobacterium tuberculosis"  
/isolate="1417-97"  
/db\_xref="taxon:1773"

gene

1. .610

/gene="rpoB"

<1. .>610

/gene="rpoB"

/codon\_start=1

/transl\_table=11

/product="RNA polymerase beta subunit"

/protein\_id="CAC87036.1"

/db\_xref="GI:2208415"

/translation="LDIYRKLPGPEPTKESAOITLLENLFFKERYDLARVRYKVNK  
KLGHVGEPTSTLTEDVATIEYLVRLHGGQTTMTVPGVEVPVDDIDHFGNR  
RLRTVGEIQNIQIRVGMRSRMRVRRMTTQDVEAITPOTLINIRPVVAAIKEFFGTS  
OPSQFMQNNPLSGLTHKRLSALGPGGLSRERAGLEVDRVHP"

BASE COUNT 122 a 191 c 202 g 95 t

ORIGIN

Query Match 58.3%; Score 124.8; DB 1; Length 610;

Best Local Similarity 77.6%; Pred. No. 4.9e-12;

Matches 166; Conservative 0; Mismatches 42; Indels 6; Gaps 1;

QY 1 TCAAGGAGAGCGCTACGACCTGGCCGGTGGCCGGGTACAAGGTGAACAAGAGCTGG 60

Db 80 TCAAGGAGAGCGCTACGACCTGGCCGGTGGCCGGGTACAAGGTGAACAAGAGCTGG 139

QY 61 GTCTTGGCGGTGCCAACCCGGCTCTGGTACTGCCACACGCTCACCGAGGAGAGCTCG 120

Db 140 GGCTGCATGTCGGCGAGCC-----CATCACGCTCGTGCAGCTGACCGAAGAAGAGCTCG 193

QY 121 TCGCCACCATCGGGTACCTGGTGGCTGCACGAGGCGCAGACCATGACCGCCCGG 180

Db 194 TGGCCACCATCGAATATCTGGTCCGCTTGACGAGGCTGACCGAAGAAGAGCTCG 253

QY 181 GCGGCTCGAGGTCCCGGTGCGAGGTGCGACGACAT 214

Db 254 GCGGCTCGAGGTCCCGGTGCGAGGTGCGACGACAT 287

RESULT 4

MTU318813

LOCUS

DEFINITION

Mycobacterium tuberculosis partial rpoB gene for RNA polymerase

beta subunit, isolate 1763-97.

ACCESSION

AJ318813

VERSION

AJ318813.1 GI:22208402

KEYWORDS

RNA polymerase beta subunit; rpoB gene.

SOURCE

Mycobacterium tuberculosis.

ORGANISM

Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;

Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium

tuberculosis complex.

REFERENCE

1

Herrera, L., Jimenez, M.S. and Saez, J.A.

Molecular analysis of rifampin-resistant Mycobacterium tuberculosis

isolated in Spain (1996-2001). Description of new alleles into rpoB

gene and review

JOURNAL

REFERENCE

2 (bases 1 to 610)

Herrera, L.

Direct Submission

Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

Unpublished

2 (bases 1 to 616)

Herrera, L.

Direct Submission

Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro

Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5,

Majadahonda. Madrid. 28220, SPAIN

FEATURES

source

1. .616

Location/Qualifiers

/organism="Mycobacterium tuberculosis"

/isolate="1763-97"

/db\_xref="taxon:1773"

1. .615

/gene="rpoB"

<1. .>615

/gene="rpoB"

/codon\_start=1

/transl\_table=11

/product="RNA polymerase beta subunit"

/protein\_id="CAC87030.1"

/db\_xref="GI:22208403"

/translation="LDIYRKLPGPEPTKESAOITLLENLFFKERYDLARVRYKVNK

KLGHVGEPTSTLTEDVATIEYLVRLHGGQTTMTVPGVEVPVDDIDHFGNR

RLRTVGEIQNIQIRVGMRSRMRVRRMTTQDVEAITPOTLINIRPVVAAIKEFFGTS

QLSQPIFMDQNNPLSGLTHKRLSALGPGGLSRERAGLEVDRVHP"

BASE COUNT 125 a 191 c 201 g 99 t

ORIGIN

Query Match 58.3%; Score 124.8; DB 1; Length 616;

Best Local Similarity 77.6%; Pred. No. 4.9e-12;

Matches 166; Conservative 0; Mismatches 42; Indels 6; Gaps 1;

QY 1 TCAAGGAGAGCGCTACGACCTGGCCGGTGGCCGGGTACAAGGTGAACAAGAGCTGG 60

Db 80 TCAAGGAGAGCGCTACGACCTGGCCGGTGGCCGGGTACAAGGTGAACAAGAGCTGG 139

QY 61 GTCTTGGCGGTGCCAACCCGGCTCTGGTACTGCCACACGCTCACCGAGGAGAGCTCG 120

Db 140 GGCTGCATGTCGGCGAGCC-----CATCACGCTCGTGCAGCTGACCGAAGAAGAGCTCG 193

QY 121 TCGCCACCATCGGGTACCTGGTGGCTGCACGAGGCGCAGACCATGACCGCCCGG 180

Db 194 TGGCCACCATCGAATATCTGGTCCGCTTGACGAGGCTGACCGAAGAAGAGCTCG 253

QY 181 GCGGCTCGAGGTCCCGGTGCGAGGTGCGACGACAT 214

Db 254 GCGGCTCGAGGTCCCGGTGCGAGGTGCGACGACAT 287

RESULT 5

MTU318815

LOCUS

DEFINITION

Mycobacterium tuberculosis partial rpoB gene for RNA polymerase

beta subunit, isolate 2540-97.

ACCESSION

AJ318815

VERSION

AJ318815.1 GI:22208406

KEYWORDS

RNA polymerase beta subunit; rpoB gene.

SOURCE

Mycobacterium tuberculosis.

ORGANISM

Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;

Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium

tuberculosis complex.

REFERENCE

1

Herrera, L., Jimenez, M.S. and Saez, J.A.

Molecular analysis of rifampin-resistant Mycobacterium tuberculosis

isolated in Spain (1996-2001). Description of new alleles into rpoB

gene and review

JOURNAL

REFERENCE

2 (bases 1 to 618)

Herrera, L.

Direct Submission

Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro

Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5,  
Majadahonda. Madrid. 28220, SPAIN

FEATURES  
source  
1..618  
/organism="Mycobacterium tuberculosis"  
/isolate="2540-97"  
/db\_xref="taxon:1773"  
1..618  
/gene="rpoB"  
<1..>618  
/gene="rpoB"  
/codon\_start=1  
/product="RNA polymerase beta subunit"  
/translation="TDEALLDIYRKLRPEPTPKESAOTLLENLFKKRYDLARVGR  
YKVNKGLHVGEPITTSLTLEEDVATIEYLRLHEGOTTMTVPGGVEVPVETDDID  
HFGNRLRTVGEILQNIQIVGMSRMERVRMTQDVEAITPTLINIRPVVAAIKE  
FFGTQSLSQFMQDNPLSLGTHKRLALPGGLSRERAGLEVRDV"  
BASE COUNT 124 a 190 c 207 g 97 t  
ORIGIN

Query Match 58.3%; Score 124.8; DB 1; Length 618;  
Best Local Similarity 77.6%; Pred. No. 4.9e-12;  
Matches 166; Conservative 0; Mismatches 42; Indels 6; Gaps 1;

QY 1 TCAAGGAGAGCGCTACGACCTGGCCCGCTGCGCGGTACAAAGTGAACGAAGCTGG 60  
|||||  
Db 95 TCAAGGAGAGCGCTACGACCTGGCCCGCTGCGCGGTACAAAGTGAACGAAGCTGG 154  
QY 61 GTCTTGGCGGTGCCAACCCGGCTCTGGTGACTGCCACGAGGAGGAGCGTGG 120  
|||||  
Db 155 GCCTGCATCTCGCGGAGCC-----CATCAGCTGCTGACGCTGACCGAAGAGCTGG 208  
QY 121 TCGCCACCATCGGTACCTGGTGGCTGCACGAGGCGGACGACGATGACCGCCCGG 180  
|||||  
Db 209 TGGCCACCATCGAATATCTGGTCCGCTTGACGAGGCTGACGACGATGACCGCTCCGG 268  
QY 181 GCGGCTCGAGTCCGCGTGGAGTGCAGACAT 214  
|||||  
Db 269 GCGGCTCGAGTCCGCGTGGAAACCGAGACAT 302

RESULT 6  
LOCUS MTU318817 618 bp DNA circular BCT 09-AUG-2002  
DEFINITION Mycobacterium tuberculosis partial rpoB gene for RNA polymerase  
beta subunit, isolate 2348-98.  
ACCESSION AJ318817  
VERSION AJ318817.1 GI:22208410  
KEYWORDS RNA polymerase beta subunit; rpoB gene.  
SOURCE Mycobacterium tuberculosis.  
ORGANISM Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium  
tuberculosis complex.  
1  
REFERENCE  
AUTHORS Herrera, L., Jimenez, M.S. and Saez, J.A.  
TITLE Molecular analysis of rifampin-resistant Mycobacterium tuberculosis  
isolated in Spain (1996-2001). Description of new alleles into rpoB  
gene and review  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 618)  
AUTHORS Herrera, L.  
TITLE Direct Submission  
JOURNAL Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro  
Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5,  
Majadahonda. Madrid. 28220, SPAIN

FEATURES  
source  
1..618  
/organism="Mycobacterium tuberculosis"  
/isolate="2348-98"

/db\_xref="taxon:1773"  
1..618  
/gene="rpoB"  
<1..>618  
/gene="rpoB"  
/codon\_start=1  
/transl\_table=11  
/product="RNA polymerase beta subunit"  
/protein\_id="CAC87034.1"  
/db\_xref="GI:22208411"  
/translation="TDEALLDIYRKLRPEPTPKESAOTLLENLFKKRYDLARVGR  
YKVNKGLHVGEPITTSLTLEEDVATIEYLRLHEGOTTMTVPGGVEVPVETDDID  
HFGNRLRTVGEILQNIQIVGMSRMERVRMTQDVEAITPTLINIRPVVAAIKE  
FFGTQSLSQFMQDNPLSLGTHKRLALPGGLSRERAGLEVRDV"  
BASE COUNT 124 a 192 c 207 g 95 t  
ORIGIN

Query Match 58.3%; Score 124.8; DB 1; Length 618;  
Best Local Similarity 77.6%; Pred. No. 4.9e-12;  
Matches 166; Conservative 0; Mismatches 42; Indels 6; Gaps 1;

QY 1 TCAAGGAGAGCGCTACGACCTGGCCCGCTGCGCGGTACAAAGTGAACGAAGCTGG 60  
|||||  
Db 95 TCAAGGAGAGCGCTACGACCTGGCCCGCTGCGCGGTACAAAGTGAACGAAGCTGG 154  
QY 61 GTCTTGGCGGTGCCAACCCGGCTCTGGTGACTGCCACGAGGAGGAGCGTGG 120  
|||||  
Db 155 GCCTGCATCTCGCGGAGCC-----CATCAGCTGCTGACGCTGACCGAAGAGCTGG 208  
QY 121 TCGCCACCATCGGTACCTGGTGGCTGCACGAGGCGGACGACGATGACCGCCCGG 180  
|||||  
Db 209 TGGCCACCATCGAATATCTGGTCCGCTTGACGAGGCTGACGACGATGACCGCTCCGG 268  
QY 181 GCGGCTCGAGTCCGCGTGGAGTGCAGACAT 214  
|||||  
Db 269 GCGGCTCGAGTCCGCGTGGAAACCGAGACAT 302

RESULT 7  
LOCUS MTU318814 633 bp DNA circular BCT 09-AUG-2002  
DEFINITION Mycobacterium tuberculosis partial rpoB gene for RNA polymerase  
beta subunit, isolate 1058-97.  
ACCESSION AJ318814  
VERSION AJ318814.1 GI:22208404  
KEYWORDS RNA polymerase beta subunit; rpoB gene.  
SOURCE Mycobacterium tuberculosis.  
ORGANISM Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium  
tuberculosis complex.  
1  
REFERENCE  
AUTHORS Herrera, L., Jimenez, M.S. and Saez, J.A.  
TITLE Molecular analysis of rifampin-resistant Mycobacterium tuberculosis  
isolated in Spain (1996-2001). Description of new alleles into rpoB  
gene and review  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 633)  
AUTHORS Herrera, L.  
TITLE Direct Submission  
JOURNAL Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro  
Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5,  
Majadahonda. Madrid. 28220, SPAIN

FEATURES  
source  
1..633  
/organism="Mycobacterium tuberculosis"  
/isolate="1058-97"  
/db\_xref="taxon:1773"  
1..633  
/gene="rpoB"  
<1..>633  
/gene="rpoB"  
/codon\_start=1

	/transl_table=1		
	/product="RNA polymerase beta subunit"		
	/protein_id="CAC87031.1"		
	/db_xref="GI:22208405"		
	/translation="DNTVGPDEALLDIYKRLRPGSPTKESAOTLLENLFKEKKRYDL ARVGRIYKNKGLGHLHGEPTSTLTEDVDVATIEYLRLHEGGOTTMTVPGGVEPVE TDDIDHFCNRLRTVCLELTONOIRVCMSEMRVVERMTTDQVEAITPTQLINIRPV AAIKEFCTSQLSOPMDONNPISGLTIYKRLSAUGPGGLSRERAGLEVDRV"		
BASE COUNT	129 a 195 c 210 g 99 t		
ORIGIN			
	Query Match 58.3%; Score 124.8; DB 1; Length 633; Best Local Similarity 77.6%; Pred. No. 4.9e-12; Matches 166; Conservative 0; Mismatches 42; Indels 6; Gaps 1;		
Qy	1 TCAAGCAGAAGCGCTACCACTGCCCGCGGTGGCCGGTACAAGGTGAACAAGAGCTGG 60       Db	110 TCAAGCAGAAGCGCTACCACTGCCCGCGGTGGCTAAGGTCAACAAGAAGCTCG 169 	
Qy	61 GTCTTGGCGGTGCCAACCCGGCTCTGTTACTGCCACCACGCTACCAGGAGNAAGCTCG 120 	Db	170 GGCTGCATGTCGGCGAGC-----CATCACGCTGCTGCAGCTGACCCAAGAAGACGCTCG 223 
Qy	121 TGCCACCATCGGTACTGTCGTGGCGCTGCACGAGGCCAGACACGATGACCGCCCCCG 180 	Db	224 TGCCACCATCAATATATGTCGGCTTGCCAGGGTGCAGACACGATGACCGTTCGG 283 
Qy	181 GCGGCTCGAGTCCCGTCCAGTCCACACAT 214 	Db	284 GCGGCTCGAGTCCCGTGGAAACCGACGACAT 317 
RESULT 8			
MTU318816			
LOCUS	MTU318816 637 bp DNA circular BCT 09-AUG-2002		
DEFINITION	Mycobacterium tuberculosis partial rpoB gene for RNA polymerase		
ACCESSION	AJ318816		
VERSION	AJ318816.1 GI:22208408		
KEYWORDS	RNA polymerase beta subunit; rpoB gene.		
SOURCE	Mycobacterium tuberculosis.		
ORGANISM	Mycobacterium tuberculosis. Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium tuberculosis complex.		
REFERENCE	Herrera,L., Jimenez,M.S. and Saez,J.A. Molecular analysis of rifampin-resistant Mycobacterium tuberculosis isolated in Spain (1996-2001). Description of new alleles into rpoB gene and review Unpublished 2 (bases 1 to 637)		
JOURNAL	Herrera,L. Direct Submission Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro National Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5, Majadahonda, Madrid 28220, SPAIN		
REFERENCE	Location/Qualifiers 1..637 /organism="Mycobacterium tuberculosis" /isolate="1255-98" /db_xref="taxon:1773" 1..637 /gene="rpoB" <1..>637 /codon_start=1 /transl_table=1 /product="RNA polymerase beta subunit" /protein_id="CAC87033.1" /db_xref="GI:22208405" /translation="VGIDEALLDIYKRLRPGSPTKESAOTLLENLFKEKKRYDLARV GRYIKYNKLGHLHGEPTSTLTEDVDVATIEYLRLHEGGOTTMTVPGGVEPVEDD		
FEATURES	source		
gene			
CDS			

Best Local Similarity		77.6%	Pred. No. 4.9e-12;	Matches 166; Conservative 0; Mismatches 42; Indels 6; Gaps 1;	
QY	1	TCAAGGAGAAGCGCTACGACCTGGCCCGCGTGGCCGGTACAAGGTGAACAAGAAGCTGG	60		
Db	101	TCAAGGAGAAGCGCTACGACCTGGCCCGCGTGGCTCGCTATAAGGTCAACAAGAAGCTCG	160		
QY	61	GTCTTGGCGGTGCCAACCCGCGCTGTGTGACTGCGACACCGCTACCGAGGAAGCTCG	120		
Db	161	GGCTGCATCTCGCGAGCC-----CATCACGCTGCTGAGGCTACCGAAGAAGACGCTG	214		
QY	121	TGCCCACCATCGGTTACCTGTGGCTGTGACGAGGCCAGACCATGATGACCGCCCGG	180		
Db	215	TGCCCACCATCGAATATCTGTGCGCTTGCACGAGGTCAGACCACGATGACCGTTCCGG	274		
QY	181	CGGGCTCGAGGTCCCGGTTCGAGGTCGACGACAT	214		
Db	275	CGGGCTCGAGGTCCGGTGAACACGACGACAT	308		
RESULT	10				
LOCUS	I50706	I50706	970 bp	DNA	linear PAT 07-OCT-1997
DEFINITION	Sequence 1 from patent US 5643723.				
ACCESSION	I50706				
VERSION	I50706.1	GI:2472409			
KEYWORDS					
SOURCE	Unknown.				
ORGANISM	Unclassified.				
REFERENCE	1 (bases 1 to 970)				
AUTHORS	Persing, D.H., Hunt, J.J., Young, K.K.Y., Felmlee, T.A., Roberts, G.D. and Whelan, A. Christian.				
TITLE	Detection of a genetic locus encoding resistance to rifampin in mycobacterial cultures and in clinical specimens				
JOURNAL	Patent: US 5643723-A 1 01-JUL-1997;				
FEATURES	Location/Qualifiers				
source	1..970				
BASE COUNT	182 a 302 c 330 g 156 t				
ORIGIN					
Query Match	58.3%; Score 124.8; DB 6; Length 970;				
Best Local Similarity	77.6%; Pred. No. 4.6e-12;				
Matches 166; Conservative 0; Mismatches 42; Indels 6; Gaps 1;					
QY	1	TCAAGGAGAAGCGCTACGACCTGGCCCGCGTGGCCGGTACAAGGTGAACAAGAAGCTGG	60		
Db	26	TCAAGGAGAAGCGCTACGACCTGGCCCGCGTGGCTCGCTATAAGGTCAACAAGAAGCTCG	85		
QY	61	GTCTTGGCGGTGCCAACCCGCGCTGTGTGACTGCGACACCGCTACCGAGGAAGACGCTCG	120		
Db	86	GGCTGCATCTCGCGAGCC-----CATCACGCTGCTGAGGCTACCGAAGAAGACGCTG	139		
QY	121	TGCCCACCATCGGTTACCTGTGGCTGTGACGAGGCCAGACCATGATGACCGCCCGG	180		
Db	140	TGCCCACCATCGAATATCTGTGCGCTTGCACGAGGTCAGACCACGATGACCGTTCCGG	199		
QY	181	CGGGCTCGAGGTCCCGGTTCGAGGTCGACGACAT	214		
Db	200	CGGGCTCGAGGTCCGGTGAACACGACGACAT	233		
RESULT	11				
LOCUS	AX111339	AX111339	3534 bp	DNA	linear PAT 30-APR-2001
DEFINITION	Sequence 2072 from Patent W00123604.				
ACCESSION	AX111339				
VERSION	AX111339.1	GI:13927631			
KEYWORDS					
SOURCE	Mycobacterium tuberculosis.				
ORGANISM	Mycobacterium tuberculosis Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;				

/product="RNA-polymerase beta subunit"	
/protein_id="AAA20242.2"	
/db_xref="GI:7144499"	
/translation="MLEGILADRSQSKTAASPSRPOSSSSNNNSVPCAPNRVSAFKL	
MSLSEDPDRFDVKAQVDECKDKMTYAAPLFTVAEFINNNTEIKSOTVPMGDFPMH	
TEKGTPIINGTERRVVSQIVRSPGVYFDETDKSDTKLHSVKVIPSRGAWLEDFVDK	
RDVGVIRDKRRQPVTVLLKALGWTSEQIVERFGFSEIMRSTLEKONTGVTDEALLD	
IYKLRPGEPPTKESAOITLLENLFKEKRYDLARVGRVKNKGLGLHVGEPITSTLIT	
EEDVATIEYLRHLEGQTITVTPGVGVVPEVETDDIDHFGNRLRTVCELTONQIRVG	
MSRNRVERMTTQDVEAITQTOTLINIRPVYAAAKEFGTSQISOFPDQNNPLSGLT	
HKRLSALPGGLSRERAGLEVDPVSHYGRMCPIETPEGPNIGLIGLSLSYARVNP	
FGIETPYKVDVGVSDIIVLTADAEDRHVYAAQANSPIDADGRFVPEPVLVRKAG	
EVEYVPSSEVDVSPROMQSVATAMIPFLEHDDANRALMGANNORQAQVPLVRSEAP	
LVGTGMLRAAIDAAATSSQESGVIEVSADYITVMHONGTRRTYMRKFAFARNHGT	
VEEDVLTSHIEHEIDARDTKLGAEEITRDIPNISDEVLDLDERGIVRIGAEVRDG	
DILGVKTPKGETELTPERLLRAJFGEKAREVRDTSKLVPHGESGKVIIGIRVESRED	
EDELPAQVNELVRVYVAQRKISDGDKLAGRNGKGVIGKILPVEDMFLADGTPVDI	
ILNTHGVPRRMNTGQILETHLGCASHGHWKVAAGKGVDPDWAARLPDELLERQPNATIS	
TPVFDGAQAEALQGLSCTLPNDRDGLVDADGKAMLFDRSGSEFPYPVTVGYMYIM	
KLHLVDDKIHARSTGYSMTITQOPLGKAQFGQGRFGEKMECWAQYAAATYLOELL	
TIKS"	
BASE COUNT	723 a 1173 c 1293 g 664 t
ORIGIN	
Query Match	58.3%; Score 124.8; DB 1; Length 3853;
Best Local Similarity	77.6%; Pred. No. 3.9e-12;
Matches 166; Conservative	0; Mismatches 42; Indels 6; Gaps 1;
QY	1 TCAAGGAGAAGCGCTACACCTGGCCCGCGTGGCCGGTACAAGTCAACAAAGAAGCTGG 60
Db	1477 TCAAGGAGAAGCGCTACACCTGGCCCGCGTGGCCGGTACAAGTCAACAAAGAAGCTGG 1536
QY	61 GTCCTGGCGGTGCCAACCGCGCTCTGTGTCGTCACACCGCTCACCGAGGAAGACGTCG 120
Db	1537 GCGTCGATGTCGGCGAGCC-----CATCAGCTGCTGCACGCTCACCGAAGAACGTCG 1590
QY	121 TGCCACCATCGGTACTGTGTCGCCCTGCACGAGGGCCAGCACCATGATGACCGCCCGG 180
Db	1591 TGCCACCATCAATATCTGTCGCCCTTGACAGAGGGTCAGACCATGATGACCGCTCCGG 1650
QY	181 GCGGCTCGAGTCCGCGTCGAGGTCGACGACAT 214
Db	1651 GCGGCTCGAGTCCGCGTGGAAACCGACGACAT 1684
RESULT 13	
MSGRPOB	5084 bp DNA linear BCT 13-SEP-1994
LOCUS	Mycobacterium tuberculosis RNA polymerase beta-subunit (rpoB)
DEFINITION	gene, complete cds and RNA polymerase beta'-subunit rpoC gene, partial cds.
ACCESSION	L27989
VERSION	1 GI:468333
KEYWORDS	RNA polymerase beta-subunit; rpoB gene.
SOURCE	Mycobacterium tuberculosis (strain Rv)
ORGANISM	Mycobacterium tuberculosis
REFERENCE	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium tuberculosis complex.
AUTHORS	Miller, L.P., Crawford, J.T. and Shinnick, T.M.
TITLE	The rpoB gene of Mycobacterium tuberculosis
JOURNAL	Antimicrob. Agents Chemother. 38 (4), 805-811 (1994)
MEDLINE	94304130
PUBMED	8031050
FEATURES	Location/Qualifiers
source	1..5084
	/organism="Mycobacterium tuberculosis"
	/strain="Rv"
	/db_xref="taxon:1773"
	1065..4598
gene	

/gene="rpoB"	
1065..4598	
/gene="rpoB"	
/codon_start=1	
/evidence=experimental	
/transl_table=11	
/product="RNA polymerase beta-subunit"	
/protein_id="AA21416.1"	
/db_xref="GI:468334"	
/translation="MLEGILADRSQSKTAASPSRPOSSSSNNNSVPCAPNRVSAFKL	
RPLEVPGLLDQTDQFEMWIGSPWRSEAAERGVNVPVGLLEEVLELSPIEDFSGS	
MSLSEDPDRFDVKAQVDECKDKMTYAAPLFTVAEFINNNTEIKSOTVPMGDFPMH	
TEKGTPIINGTERRVVSQIVRSPGVYFDETDKSDTKLHSVKVIPSRGAWLEDFVDK	
RDVGVIRDKRRQPVTVLLKALGWTSEQIVERFGFSEIMRSTLEKONTGVTDEALLD	
IYKLRPGEPPTKESAOITLLENLFKEKRYDLARVGRVKNKGLGLHVGEPITSTLIT	
EEDVATIEYLRHLEGQTITVTPGVGVVPEVETDDIDHFGNRLRTVCELTONQIRVG	
MSRNRVERMTTQDVEAITQTOTLINIRPVYAAAKEFGTSQISOFPDQNNPLSGLT	
HKRLSALPGGLSRERAGLEVDPVSHYGRMCPIETPEGPNIGLIGLSLSYARVNP	
FGIETPYKVDVGVSDIIVLTADAEDRHVYAAQANSPIDADGRFVPEPVLVRKAG	
EVEYVPSSEVDVSPROMQSVATAMIPFLEHDDANRALMGANNORQAQVPLVRSEAP	
LVGTGMLRAAIDAAATSSQESGVIEVSADYITVMHONGTRRTYMRKFAFARNHGT	
ANOCPLVDAGRVEAGOVADGPTDGDGEMALGNLLVAIMPWEHNYEDAIILSNRL	
VEEDVLTSHIEHEIDARDTKLGAEEITRDIPNISDEVLDLDERGIVRIGAEVRDG	
DILGVKTPKGETELTPERLLRAJFGEKAREVRDTSKLVPHGESGKVIIGIRVESRED	
EDELPAQVNELVRVYVAQRKISDGDKLAGRNGKGVIGKILPVEDMFLADGTPVDI	
ILNTHGVPRRMNTGQILETHLGCASHGHWKVAAGKGVDPDWAARLPDELLERQPNATIS	
TPVFDGAQAEALQGLSCTLPNDRDGLVDADGKAMLFDRSGSEFPYPVTVGYMYIM	
KLHLVDDKIHARSTGYSMTITQOPLGKAQFGQGRFGEKMECWAQYAAATYLOELL	
TIKSDDTVGRVKYEAIVKGENIPGPIPESEKVLKLEQLSLNEVLSDDGAAIEL	
REGDEDELERAAANUGINLSRNESAFEDLA"	
4641..5084	
/gene="rpoC"	
4641..5084	
/gene="rpoC"	
/codon_start=1	
/transl_table=11	
/product="RNA polymerase beta'-subunit"	
/protein_id="AAA21417.1"	
/db_xref="GI:537608"	
/translation="MLDVNFFDELRLGLATAEDIRQMSYGEVKKKPTINRYLKPDK	
GLFCEKIFGPTDRWECYCKGKVRKFKGIICERCQVEVTRAKVRRMRMGHIAAPVT	
HIWFKGVPSRLGVLDDLPKLEKIIFYAAYVITSVDEMRHNEI"	
BASE COUNT	969 a 1534 c 1691 g 890 t
ORIGIN	
Query Match	58.3%; Score 124.8; DB 1; Length 5084;
Best Local Similarity	77.6%; Pred. No. 3.7e-12;
Matches 166; Conservative	0; Mismatches 42; Indels 6; Gaps 1;
QY	1 TCAAGGAGAAGCGCTACACCTGGCCCGCGTGGCCGGTACAAGTGAACAAAGAAGCTGG 60
Db	1966 TCAAGGAGAAGCGCTACACCTGGCCCGCGTGGCTGCTATAAGGTCAACAAAGAAGCTCG 2025
QY	61 GTCCTGGCGGTGCCAACCGCGCTCTGTCGACCTGCCACCCAGCTCACCGAGGAAGACGTCG 120
Db	2026 GCGTCGATGTCGGCGAGCC-----CATCAGCTGCTCGAGCTGACCGAAGAACGTCG 2079
QY	121 TGCCACCATCGGTACTGTGTCGCCCTGCACGAGGGCCAGCACCATGATGACCGCCCGG 180
Db	2080 TGCCACCATCGAATATCTGTCGCCCTTGACAGAGGGTCAGACCATGATGACCGCTCCGG 2139
QY	181 GCGGCTCGAGTCCGCGTCGAGGTCGACGACAT 214
Db	2140 GCGGCTCGAGGTCGCGGTGGAAACCGACGACAT 2173
RESULT 14	
AE006964	19352 bp DNA linear BCT 27-APR-2001
LOCUS	Mycobacterium tuberculosis CDC1551, section 50 of 280 of the
DEFINITION	complete genome.
ACCESSION	AE006964 AE000516
VERSION	AE006964.1 GI:13880217

KEYWORDS	Mycobacterium tuberculosis CDC1551.
SOURCE	Mycobacterium tuberculosis CDC1551
ORGANISM	Bacteria; Firmicutes; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium tuberculosis complex.
REFERENCE	1. (bases 1 to 19352)
AUTHORS	Fleischmann, R.D., Alland, D., Eisen, J.A., Carpenter, L., White, O., Peterson, J., Deboy, R., Dodson, R., Gwinn, M., Haft, D., Hickey, E., Kolonay, J.F., Nelson, W.C., Umayam, L.A., Ermolaeva, M., Salzberg, S.L., Delcher, A., Utterback, T., Weidman, J., Khouri, H., Gill, J., Mikula, A. and Bishai, W.
TITLE	Whole genome comparison of Mycobacterium tuberculosis clinical and laboratory strains
JOURNAL	Unpublished
REFERENCE	2. (bases 1 to 19352)
AUTHORS	Fleischmann, R.D., Alland, D., Eisen, J.A., Carpenter, L., White, O., Peterson, J., Deboy, R., Dodson, R., Gwinn, M., Haft, D., Hickey, E., Kolonay, J.F., Nelson, W.C., Umayam, L.A., Ermolaeva, M., Salzberg, S.L., Delcher, A., Utterback, T., Weidman, J., Khouri, H., Gill, J., Mikula, A. and Bishai, W.
TITLE	Direct Submission
JOURNAL	Submitted (25-APR-2001) The Institute for Genomic Research, 9712
FEATURES	Medical Center Dr, Rockville, MD 20850, USA
source	Location/Qualifiers
	1. .19352
	/organism="Mycobacterium tuberculosis CDC1551"
	/strain="CDC1551"
	/db_xref="taxon:8331"
gene	/note="clinical strain"
	163. .3699
CDS	/gene="MT0695"
	163. .3699
	/gene="MT0695"
	/note="similar to GB:L27989 GB:L05910 GB:U12205 SP:P47766
	PID:149992; identified by sequence similarity; putative"
	/codon_start=1
	/transl_table=11
	/product="DNA-directed RNA polymerase, beta subunit"
	/protein_id="AAK44921.1"
	/db_xref="GI:13880218"
	/translation="MLEGCIADSRQSKTAASPSRQSSNNVPGAPNRYVFAKL
	REPLEVGLLDVQTFSEWLGISPRNRESAERGDNVPGLEEVLEIYELSPIDFSGS
	MSLSFGPDDVAPVDECKDMVTAAPLFVTAEPINNTSGEIKSOTFWMGDFPM
	TEKGTFTINTERVVSQVLRSPCVFDEITDKTDKTLHSVKVIPSRAGMFEEDVK
	RDYGVIRDKRRQPVYLLKALNTSEQIVREFGSEIMRSLERKDNVGTDEALLD
	IYRLRGEPTRESAQTLLNFFREKRYDLARVKYKVNKKLGLRHVGEPITSSLT
	EEDVATIEYLRHLEGOTTMTVPFGVEVPEVTDIDHFGNRLRTVGEIQQIRVG
	MSRMERVREMTQDVEAITPQTLINIRPVAAIKFEFGTSQSFQMDQNNPLVSLIT
	HKRLSALPGLSRRAGLEVDPVHPSHYRCMPTETPEGNIGLISGSLVRYNP
	EGFLETYPKRVQGVVSDIEVYLTADEEDRHHVVAQANSPIDADGRFEVPEVLRKAG
	EVEYVPSSEVDYHDVSPQWVSATANIPELHDDANRALGMANQOAVPLVRSEAP
	LVGTGMLRAIDAGDVVAEESGVEEVASDYITVMHONGTRRTYMRKFAKSNHGT
	CANQCPVADGRVEAGQVADGCTDGMALGNKLLVAIMPWEHNYEDAILNSR
	LVEEDVLTSTHIEHEIDARDTLGAEITRDPINISDEVLADLDRGIVRIGAEVRD
	GDILVGKTPKGETPEERLLRAIFGEKAREVROTSLKVPHGSGKPIGVFVSRE
	DEDELPAVNELRVYVAOKRISDQDKLAGRGNKVICKILPVEDMPPLADGTPVD
	IILNTHGVPRMNIQOILETHLGNCAHSCHKVDAAKGVPDPAARLDELLEAOPNAIV
	STPVFDGAQLGGLSCLTPNRRDVLVDADKAMLPDRSGRGEFPYPTVGYMTI
	MKHLHVDKXIHARSTGYSMTQQPLGKAQFGQGRFGEKMECNAWQAYGAAYTLQEL
	LTIKSDDTGVRKVYEAIVKGENIPEGIPESFKVLLKELQSLCLNVEVLSDDGAIE
	LREGEDEDLERAAANGINLSRNESASVEDLA"
	3744. .7694
gene	/gene="MT0696"
	3744. .7694
CDS	/gene="MT0696"
	/note="similar
	similarity; putative"
	/codon_start=1
	/transl_table=11
	/product="DNA-directed RNA polymerase, beta-prime subunit"
	/protein_id="AAK44922.1"
	/db_xref="GI:13880219"
	/translation="MLDVNFDELRIGLATAEDIRQSYGEVKKPETINYBTLKPEKD
	GLFCRIFGPTDWCYCGKTKRVRFRKGLICERCVCYETRAKVRRRRNGHLELAAPVT
	HIWYFKGPSRLGYLLDLAPLKDLEKIYIFAAVYITVSDEMRHNELSLEAMAVERK
	AVEDQGELEAARQKLELAELAEGAADARKVRDGRGERMRQIRDAQRELDAR
	LEDIWSFTFKLAPKOLIVDENLYRELVDYGEYFTGAMGAESIOLKLIENFIDAEAS
	LRDVRNGGKKRLAKRLKVAAFQSGNSMGWLDADVPIPELPRMVLQDGGRR
	FATSDNLDIRVINRNRKRLIDGLAPELIVNNEKRLQESVDALFDNGRRGPPVT
	GGGNRLPLSLLKQGRFRQNLGKRVDYSGRSVIVGFPQLKHOCCGLPKLMALE
	LFPKPMKRLVDLNAHQNKSAKRMVERFORPOVMDVLEEVIAEHFVLLNLRAPTLLRLG
	IQAFPEMLVEGKAIQHLPLVEAFNAFEDGDMVHLPLLSAEQAARILMLSSNNIL
	SPASGRPLAMPRLDMVTGLYLTTEVPEDTGEYQPSADGHPETGYVSSPAIAINAADR
	GVLSVRKIKVRLTOLRPPEVEIAELFCHSGMOPGDMAWETTIGRVMFNELLPLGYR
	FVNKOMHKKVQAAIINDLAERYPMIVVAQTVDKLKDAFYHATSEVTVSNADVLNPP
	RKKEILDHYEERADKVEKQFORGALNHDERNEALVEIKKEATDEVGQALREHYDPDNP
	IITIVDSGATGNFTOTRTLAGMKGLVTPKGEFIPRPKSSFRGLTVLEVFINTHGA
	RGLADTALTADSGYLTRRLVDVSDQVIVREHDCQTERGIVELAEARAPDGLIRDP
	YLETSAYARTLTGTDVADAGNIVVERGQDLGDPEDALLAAGITGVKRSVLTCTATST
	GVCACTYGRSMATKGLVDIGEAAGVIAAQSIGEPGCTQTMRTFGGVGDEDITGGLPR
	VOELFEARVPRGKAPIDVTVGRVLEDGEREVKITIVPDDGSEVVYDKISKRORLRV
	FKHEDGSESVLSDGCHVEVGGQLMGSDAPHEVLRVQGPREVQVHLREVQEVYRAQG
	VSIHDKHIEVIVRQMLRRVTIIDSSTFELPGLSLIDRAEFKAENRRVVAEGGEPAAGR
	PVLMGTTKASLATDSMLSAASFQETTRVLTDAAINCRSKLNGKENYIIIGKLIPAGT
	GINRYRNTAVOPTEEARAAYTIPSYEQYYSDFGAATGAAPVLDLDYYSIDYR"
	complement(7691. .8065)
gene	/gene="MT0697"
	complement(7691. .8065)
CDS	/gene="MT0697"
	/note="identified by Glimmer2; putative"
	/codon_start=1
	/transl_table=11
	/product="hypothetical protein"
	/protein_id="AAK44923.1"
	/db_xref="GI:13880220"
	/translation="MFDAAAITNPGHAWASAMERSGLLECVAGLEQPFGEFTADKL
	NPDGSSRRVPRRQADGGIATHVERGGQSGQGVVPMHGFALAMQDRLIHH
	GGQTNQRIAAQFVRVFCVCSPT"
	complement(8058. .9972)
gene	/gene="MT0698"
	/note="this region contains an authentic frame shift and
	is not the result of a sequencing artifact; identified by
	Glimmer2; putative; conserved hypothetical protein,
	authentic frameshift"
	10167. .10925
gene	/gene="MT0699"
	10167. .10925
CDS	/gene="MT0699"
	/note="identified by match to PFAM protein family HMM
	PF011261"
	/codon_start=1
	/transl_table=11
	/product="AP endonuclease, family 2"
	/protein_id="AAK44924.1"
	/db_xref="GI:13880221"
	/translation="MLIGSHVSPDTPLAAAEEAGADVQVIFLGNPQSWKAPKPRDDAA
	AKAATLPIYHAPYLINLASANNRVRIPSRKILOETCAAAADICAAAVIVHGHHVAD
	DNDIDKGQWRKALDRLETEVPVYLENTAGGDHAKAREDTIARLMDVIGDTGIGFC
	LDTCHTWAGEALTDVADRIKAITCRIDLVLHCNDSRDEANGSRDRHANLGSQIDPDL
	LVAAKTAAGPVCITADQGRKDDIAFLRERTGS"
	10957. .11799
gene	/gene="MT0700"
	10957. .11799
CDS	/gene="MT0700"
	/note="similar to GB:U00012 PID:466863; identified by
	sequence similarity; putative"
	/codon_start=1
	/transl_table=11
	/product="hydrolase/esterase, putative"
	/protein_id="AAK44925.1"
	/db_xref="GI:13880222"
	/translation="MLRRVAILLAVALFAGCGGTRLAAGFGNSVHTLDVGDAGR
	SVRLYKPGCLSPSSAPLVLMHGGFGSAKQERSYGMDELASEKLELVAYIVHGHRWN
	ANGGGCCRPAREGVDDIGFVRVYADIANVYSIDPARVYVTMGNGAIIMSYTLACNT
	SIFAAIGVYSGTQLDPCQSPRPVSVIHITGTADPLVRVHGGPGAGFARIDGPPVVDLN



```
AFWREVNRCGALDTTTEGVPVTTSGATCADNRVRVLLTVDDAGHRWPSFATQTLNRFFA
AHER"
11859..13487
/genes="MT0701"
11859..13487
/genes="MT0701"
/notes="similar to SP:P33224 GB:L20915 PID:457172
PID:457174 PID:537028; identified by sequence similarity;
putative"
/codon_start=1
/transl_table=11
/product="acyl-CoA dehydrogenase, putative"
/protein_id="AAK44926.1"
/db_xref="GI:13880223"
/translation="MSDTHVTVNQVPLENYNPASPVLIEALIOEGGOMGLDEVNEY
GAISACQQRGEALADNRPIILHTDAYVDEVEYDPAYHELMRTAITHGMHAAP
WADRRGAGGAEALADNRPIILHTDAYVDEVEYDPAYHELMRTAITHGMHAAP
DELPKATTPKAGITAGMSMTEKGGSDVRAGTTPADNAGSYSLTHGKWTSAVPMCD
IFLVLAQAPDGLSCFLPRVLPDGTNRNMFQRLDKLGNHANASSEVEYDGAVALV
GEGRCVPTLIEVNLTRLDLCALGSATSMRTGLTRAVHHAQHRKAFAYLIDQPLRN
VLADLVEAEATIVAMRNAGATDNARVGNTEALLRIGLAAAYWCKRSTAHAAE
ALECGNGYVSDSGMPLRYREAPLMGWEGSGNVSALDTLRAMATRPACVEVLFDEL
ARSAQDPLRDLRVRPQLGDLDTIGYRARKIAEDICLALQGLSLVRVHGHPAVAEA
FLATRLGGQMGAYGTMPPAGLDLAPILERALVKG"
13498..14436
/genes="MT0702"
13498..14436
/genes="MT0702"
/notes="similar to GP:3885480; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="enoyl-CoA hydratase/isomerase family protein"
/protein_id="AAK44927.1"
/db_xref="GI:13880224"
/translation="MTHAIRPVDLNTMTYEVTRIGRIATFNRPKGNALIIDPTPL
ELSALVERADLPQVHVILVSGRGFCAGFDLSAAGSSSTGGGAYQGVLDGKT
QAVNHLNPQWPMIDYQNMRSFVRGFASLMHADRTYVKIHGYCVAGGTDLALHADQ
VIANADAKGYPTPRVWGPVAGLWHLRQDQRAKRLFTFCITCAGAAEMGLAVEA
PEGADVATERLVAALPYNQLIMVKLANSALLOOGVATSRMSTVDFGAAHRT
PGHAFADAVERGFDAVRRDEPFGDYGROASRV"
14439..15161
/genes="MT0703"
14439..15161
/genes="MT0703"
/notes="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
Query Match 58.3%; Score 124.8; DB 1; Length 19352;
Best Local Similarity 77.8%; Pred. No. 3.1e-12;
Matches 166; Conservative 0; Mismatches 42; Indels 6; Gaps 1;
Qy 1 TCRAAGGAGAGCGCTACGACCTGCCCGCTGGCGGCTACAAAGGTGAACAAGACTCG 60
Db 1064 TCRAAGGAGAGCGCTACGACCTGCCCGCTGGCTGCTATAGGTCACAGAAGCTCG 1123
Qy 61 GTCTTTGGCGGTGCCAACCCGGCTCTGGTGACTGCCACACCGCTCCAGGAAGAGCTCG 120
Db 1124 GGCTGTCATCTCGCGAGGCC-----CATCACGTCGTCGACGCTGACCGAAGAGAGCTCG 1177
Qy 121 TCGCCACCATCGGTACCTGGTGGCTGCACGAGGCGCCACGACGATGACCGCCCCCG 180
Db 1178 TGGCCACCATCGGAATATCTGGTCCGTTGTCACGAGGGTCAGACCAGATGACCGTTCCCG 1237
Qy 181 GCGGCCCTCGAGGTCCCGGTCGAGGTGCACGACAT 214
Db 1238 GCGCGCTCGAGGTGCCGGTGAACACGACGACAT 1271
RESULT 15
MTCI376
```

LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

MTCI376  
Mycobacterium tuberculosis H37Rv complete genome; segment 32/162.  
Z95972.1 GI:3261790  
Mycobacterium tuberculosis H37Rv.  
Mycobacterium tuberculosis H37Rv.  
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
Actinomycetales; Corynebacterineae; Mycobacteriaceae;  
Mycobacterium; Mycobacterium tuberculosis complex.  
1 (bases 1 to 19770)  
Cole, S. T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C.,  
Harris, D., Gordon, S. V., Eiglmeier, K., Gas, S., Barry III, C. E.,  
Tekala, F., Badcock, K., Basham, D., Brown, D., Chillingworth, T.,  
Connor, R., Davies, R., Devlin, K., Feltwell, T., Gentles, S.,  
Hamlin, N., Holroyd, S., Hornsby, T., Jagels, K., Krogh, A., McLean, J.,  
Moule, S., Murphy, L., Oliver, S., Osborne, J., Quail, M. A.,  
Rajandream, M. A., Rogers, J., Rutter, S., Seeger, K., Skelton, S.,  
Squares, S., Squares, R., Sulston, J. E., Taylor, K., Whitehead, S., and  
Barrell, B. G.

REFERENCE  
AUTHORS

Deciphering the biology of Mycobacterium tuberculosis from the  
complete genome sequence  
Nature 393 (6685), 537-544 (1998)  
98295987  
9634230  
2 (bases 1 to 19770)  
Direct Submission  
Parkhill, J.

TITLE

Submitted (11-JUN-1998) Submitted on behalf of the Mycobacterium  
tuberculosis sequencing and mapping teams, Sanger Centre, Wellcome  
Trust Genome Campus, Hinxton, Cambridge CB10 1SA Unite de Genetique  
Moleculaire Bacterienne, Institut Pasteur, 28 rue du Docteur Roux,  
75724 Paris Cedex 15, France E-mail: parkhill@sanger.ac.uk  
On Jun 27, 1998 this sequence version replaced gi:2143285.

COMMENT

Notes:  
Details of M. tuberculosis sequencing at the Sanger Centre are  
available on the World Wide Web.  
(URL, [http://www.sanger.ac.uk/projects/M\\_tuberculosis/](http://www.sanger.ac.uk/projects/M_tuberculosis/)) CDS have  
been renumbered from the original cosmid submissions but the old  
gene designations are in brackets after the new gene numbers.  
Gene prediction was based on a Hidden Markov Model of TB genes  
implemented in TBparse (Krogh) supplemented with visual inspection  
of positional base preference in codons, especially where there is  
an increase in the observed/expected third position G + C.  
CAUTION: In some cases we may not have predicted the correct  
initiation codon. Where possible we choose an initiation codon  
(atg, gtg, or ttg) which is preceded by an upstream ribosome  
binding site sequence (optimally 5-13bp before the initiation  
codon). If this cannot be identified we choose the most upstream  
initiation codon.

FEATURES  
source

Location/Qualifiers  
1..19770  
/organism="Mycobacterium tuberculosis H37Rv"  
/strain="H37Rv"  
/db\_xref="taxon:83332"  
/clone="I376"  
57..61  
/note="agg, possible rbs upstream of Rv0654"  
68..1573  
/gene="Rv0654"  
68..1573  
/gene="Rv0654"  
/note="Rv0654", (MTCI376.22), len: 501. unknown, FASTA  
score: Q53353 LIGNOSTILBENE-ALPHA, BETA-DIOXYGENASE (485  
aa) opt:280 z-score: 330.1 E(): 2.3e-11, (28.5% identity  
in 523 aa overlap). Also similar to M. tuberculosis protein  
MTCV21C12.07C (29.5% identity in 522 aa overlap).  
/codon\_start=1  
/transl\_table=11  
/product="hypothetical protein Rv0654"  
/protein\_id="CAB09380.1"  
/db\_xref="GI:2143307"  
/db\_xref="SPTREMBL:O06785"

RBS

gene

CDS

	<pre>/translation="MTTAAAEASQNPYLEGLAPVSTEVATDLPVGTGRIPEHLGDRY LRGNPNVAVDPATYHWTGDMVHGVALRQCKARWYRNRYRVPVPAALAGEPISA RPHPTGIIIEGSPNTVNLVTHAGRTLALVAGVYNYELTDDELDTVPGCDFGTLHGYYT AHQDPDPTGELHNAVSTFARGHRVQISVIGDGHARTVTDIEVAGSPMHHSFSLTDN YVYIDLPVTFDPMQVVPASVPRMQLQRPALVYQSLGRVIRPDPIAALGNRMQHSO RLPYAMNPSPARVGPREGNEDVWFQIEPCVYIHPNLNAYSCRNAEVLVLDDVY RYSRMEDRRGRGDSRSLDRWTINLATGAVTAECDRDRAQEPFRINETLVGGPHR FAVTVTEGGLFVAGAAALSTPLYKDCVTSSTVASLDPDLLIGEMVFPVNPASRAE DDGILMCGYGHRCRGDEGQLLLDQAQTLESIAIVHLPRQVPMGFHGNWAPTT" 1585. 2664 /genes="Rv0655" /len="2664" /notes="Rv0655" /notes="Rv0655" FASTA score: YRBF_ECOLI P45393 hypothetical abc transporter atp-binding (269 aa), opt: 644 z-score: 721.8 E(): 3.4e-33 (38.5% identity in 244 aa overlap); contains PS00017 ATP/GTP-binding site motif A, PS00211 ABC transporters family signature, highly similar to M. leprae MKL_MYCLE P30769 possible ribonucleotide transport atp- (347 aa) opt: 2021 z-score:2244.4 E(): 0. (92.2% identity in 335 aa overlap). Also similar to many otherM. tuberculosis ABC transporters eg. MTCY253.24 (33.6% identity in 241 aa overlap)" /codon_start=1 /transl_table=11 /product="hypothetical protein Rv0655" /protein_id="CAB09379.1" /db_xref="GI:2143306" /db_xref="SWISS-PROT:O06784" /translation="MRYSDSYHTGWRQPRASTGFPMGVSIENVGLTKFGSSRIWE DVTLPAGEVSLVIGSGTSKVFSLKSLGLLRPERGSIIIDGTIDIEGSAKELVEI RTFLVGFODGALFGSMNLVNTAFPLRHTKKKESEITDIVMEKALVGLGDEKFF PEIEGGHMKRAGLALVLDPQIILCEPDSDGLDPVTVLSOLINDJNAQIDARTIL IYTHNINTARTVPDNGMLFRKHLVFMFGPREVLITSDPVVROPLNGRRIGPMSEE KDEATMAEALLDAGHAGGVEETEGVPPQISATPCMPERKAVARQARVREMLTFL PKQAQAAILLDLEGHKAYVHEIGQ" 1762. 1785 /genes="Rv0655" /notes="Rv0655" /notes="PS00017 ATP/GTP-binding site motif A" 2074. 2118 /genes="Rv0655" /notes="PS00211 ABC transporters family signature" complement(3052. .3435) /genes="Rv0656c" complement(3052. .3435) /genes="Rv0656c" /notes="Rv0656c, (MTCI376.20, unknown), len: 127 aa" /codon_start=1 /transl_table=11 /product="hypothetical protein Rv0656c" /protein_id="CAB09379.1" /db_xref="GI:2143305" /db_xref="SPTREMBL:O06783" /translation="MAAATTGTGHRGLELRAAORAVGSCPEPORAECRSARNADFPQ MSRMFGDVPDPVPKPSVWRMIDSAQHRLARAGAVGALSVDVLLICTDAARGLVVLH DDADYELAEHLRPLDIRVRVVSADD" complement(3530. .3685) /genes="Rv0657c" complement(3530. .3685) /genes="Rv0657c" /notes="Rv0657c, (MTCI376.19), unknown, len: 51 aa: similar to several other M. tuberculosis hypothetical proteins eg. YW08_MYCTU Q10848 hypothetical 8.9 kd protein cy39.08c (80 aa), fasta scores; opt: 107 z-score: 182.3 E(): 0.0036, 45.8% identity in 48 aa overlap. Also similar to MTCY48.5 andAL020958 SC4H8_7 streptomycetes coelicolor cosmid 4H8 (66 aa), 41.0% identity in 39 aa overlap." /codon_start=1 /transl_table=11 /product="hypothetical protein Rv0657c" /protein_id="CAB09377.1" /db_xref="GI:2143304"</pre>	
misc_feature		
misc_feature		
gene		
CDS		
gene		
CDS		
gene		
CDS		
Query Match	58.3%;	Score 124.8; DB 1; Length 19770;
Best Local Similarity	77.6%;	Pred. No. 3.1e-12;
Matches	166;	Conservative 0; Mismatches 42; Indels 6; Gaps 1;
Qy	1	TCAAGGAGAGCGGTACGACCTGGCCGCGGTGGCGGTGACAGGTGAACAAGAGCTGG 60
Db	10758	TCAAGGAGAGCGGTACGACCTGGCCGCGGTGGCGGTGACAGGTGAACAAGAGCTGG 10817
Qy	61	GTCTTGGCGGTGCCAACCCGGCTCTGGTACTGCCACACGCTCACGAGGAAGAGCTGG 120
Db	10818	GGCTGCATGTGGCGGAGGCC-----CATCACGCTGTCGACGCTGACCGAAGAGAGCTGG 10871
Qy	121	TGCCACCATCGGTACTGTGGCTGCGCTGCACGAGGCGCCAGACGATGACCGCCCGG 180
Db	10872	TGGCCACCATGGAATATCTGTCGCTGTCGAGGAGGTGACACGATGACCGTTCCGG 10931
Qy	181	CGGCGCTCGAGGTCCCGGTGAGGTGACGACAT 214

Db 10932 GCGCGTCGAGGTGCCCGTGGAAACCGACGACAT 10965

Search completed: November 13, 2002, 01:25:30  
Job time : 681.07 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 12, 2002, 16:59:42 ; Search time 1109.13 Seconds  
(without alignments)  
3037.202 Million cell updates/sec

Title: US-09-697-123B-1  
Perfect score: 208  
Sequence: 1 tcaaggagagcgtacgac.....ccggttgagaccgacacat 208

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues  
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*

- 1: em\_estba:\*
- 2: em\_esthum:\*
- 3: em\_estin:\*
- 4: em\_estmu:\*
- 5: em\_estov:\*
- 6: em\_estpl:\*
- 7: em\_estro:\*
- 8: em\_hic:\*
- 9: gb\_estl:\*
- 10: gb\_est2:\*
- 11: gb\_hic:\*
- 12: gb\_est3:\*
- 13: gb\_est4:\*
- 14: gb\_est5:\*
- 15: em\_estfun:\*
- 16: em\_estom:\*
- 17: gb\_gss:\*
- 18: em\_gss\_hum:\*
- 19: em\_gss\_inv:\*
- 20: em\_gss\_pln:\*
- 21: em\_gss\_vrt:\*
- 22: em\_gss\_fun:\*
- 23: em\_gss\_mam:\*
- 24: em\_gss\_mus:\*
- 25: em\_gss\_other:\*
- 26: em\_gss\_pro:\*
- 27: em\_gss\_rtd:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	163.2	78.5	1282	9	AI770311 42 Mycoba
2	47	22.6	517	13	BJ233830 BJ233830
3	47	22.6	649	12	BC908023 TAlr1164F
4	47	22.6	684	13	BJ258518 BJ258518
5	47	22.6	1013	10	BE418320 SCL024.F0
6	45.4	21.8	363	12	BC907534 TAlr1160G

7	45.2	21.7	468	14	BO463007
8	45.2	21.7	479	14	BM987859
9	45.2	21.7	747	13	BI914649
10	45.2	21.7	898	13	BI755032
11	44.8	21.5	988	10	BE213908
12	44.6	21.4	616	13	BI722737
13	44.2	21.3	686	9	AL509174
14	44.2	21.3	877	17	AO687771
15	43.8	21.1	568	13	BI995346
16	43.8	21.1	638	12	BG415787
17	43.6	21.0	421	10	BE230529
18	43.6	21.0	985	12	BG338508
19	43.4	20.9	694	13	BJ457580
20	43.2	20.8	460	13	BM100319
21	43.2	20.8	1888	11	AY104150
22	43	20.7	558	13	BM660850
23	42.8	20.6	553	13	BI682538
24	42.8	20.6	563	10	BE666613
25	42.8	20.6	623	13	BM140530
26	42.8	20.6	648	14	BO606331
27	42.8	20.6	788	10	BE413803
28	42.8	20.6	1730	13	BM479096
29	42.6	20.5	925	17	CNS0091P
30	42.4	20.4	448	12	BG463919
31	42.4	20.4	519	10	AM676807
32	42.4	20.4	539	12	BG556957
33	42.4	20.4	589	12	BG322881
34	42.4	20.4	595	10	BE357434
35	42.4	20.4	367	10	BE586511
36	42.2	20.3	367	10	BE586511
37	42.2	20.3	573	14	BQ246115
38	42.2	20.3	630	12	BG906999
39	42.2	20.3	649	14	BQ246063
40	42	20.2	357	14	BQ809506
41	42	20.2	444	17	AQ936743
42	42	20.2	710	12	BG321153
43	42	20.2	779	9	AU067540
44	42	20.2	888	9	AU051760
45	42	20.2	1357	11	BC029965

ALIGNMENTS

RESULT 1  
AI770311/c  
LOCUS AI770311 1282 bp mRNA linear EST 24-JAN-2000  
DEFINITION 42 Mycobacterium anaerobic stationary phase library Mycobacterium smegmatis cDNA, mRNA sequence.  
ACCESSION AI770311  
VERSION AI770311.1 GI:6742680  
KEYWORDS EST.  
SOURCE Mycobacterium smegmatis.  
ORGANISM Mycobacterium smegmatis  
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
REFERENCE 1 (bases 1 to 1282)  
AUTHORS Murugasu-Oei, B., Tay, A. and Dick, T.  
TITLE Upregulation of stress response genes and ABC transporters in anaerobic stationary-phase Mycobacterium smegmatis  
JOURNAL Mol. Gen. Genet. 262 (4-5), 677-682 (1999)  
MEDLINE 20092472  
COMMENT Contact: Murugasu-Oei, B.  
Mycobacterium Laboratory  
Institute of Molecular and Cell Biology  
30 Medical Drive, Singapore 117609, Republic of Singapore  
Tel: 65 874 3011  
Fax: 65 779 1117  
Email: mcbom@imcb.nus.edu.sg  
Insert Length: 1282 Std Error: 0.00  
Seq primer: T3 Forward; T7 backward.

FEATURES

source

```

/organism="Mycobacterium smegmatis"
/strain="mc2155"
/db_xref="taxon:1772"
/clone_lib="Mycobacterium anaerobic stationary phase
library"
/lab_host="E. coli XL1-Blue MRP"
/notes="Vector: Lambda ZAP II; Bacilli were disrupted using
a Mini bead beater (Biospec). RNA was isolated using the
RNeasy protocol (Qiagen). Purified RNA was subjected to 2
rounds of digestion with RNase-free DNase I (Promega).
DNase I was heat-inactivated at 75°C for 5 min. and
removed by using RNeasy columns followed by phenol
extraction and ethanol precipitation. The RNA
preparations were confirmed to be free of genomic DNA
contamination by carrying out PCR and RT-PCR using the
Access kit (Promega) and primers specific for the
histone-like protein gene hlp (Lee et al., 1998). cDNA
was synthesized using random hexamer primers (Promega)
and Stratagene's cDNA synthesis kit. cDNA fragments were
ligated into Lambda ZAP II vector and packaged in vitro
using Stratagene reagents."
BASE COUNT      230 a 395 c 405 g 252 t
ORIGIN

Query Match      78.5%; Score 163.2; DB 9; Length 1282;
Best Local Similarity 86.5%; Pred. No. 4.5e-26;
Matches 180; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

QY 1 TCAAGGAGAACGCTACGACCTGGCCGGGTAGGCGCTACAAGGTCAACAAGAAGCTCG 60
|||||
Db 236 TCAAGGAGAACGCTACGACCTGGCCGGGTAGGCGCTACAAGGTCAACAAGAAGCTCG 177
|||||

QY 61 GCCTGACGCTCGCGGATCCGATCACAGCTCCACGTCACGAGGAGACGCTCGCGCA 120
|||||
Db 176 GCCTGACGCGGCAAGCGGATCACAGCTCCACGTCACGAGGAGACGCTCGCGCA 117
|||||

QY 121 CCATGAGTACCTGCTGCGCTGACGAGGCGCCACGACGATGACCTCCCGGCGGCA 180
|||||
Db 116 CCATGAGTACCTGCTGCGCTGACGAGGCGCTGACGATGACCTCCCGGCGGCG 57
|||||

QY 181 CCGAGGTGCGGTTTCGACCGCAGACAT 208
|||||
Db 56 TCGAGTTCCCGTCGAGGTCGACGACAT 29
|||||

RESULT 2
BJ233830
LOCUS      BJ233830 517 bp mRNA linear EST 05-APR-2002
DEFINITION aestivum cDNA clone whe7g20 5', mRNA sequence.
ACCESSION  BJ233830
VERSION     BJ233830.1 GI:20050958
KEYWORDS   bread wheat.
SOURCE     Triticum aestivum
ORGANISM   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
; Triticeae; Triticum.
REFERENCE  1 (bases 1 to 517)
AUTHORS   Ogihara, Y. and Murai, K.
TITLE     Expressed genes in Triticum aestivum
JOURNAL   Unpublished (2002)
COMMENT   Contact: Tadasi Shin-i
          Center For Genetic Resource Information
          National Institute of Genetics
          1111 Yata, Mishima, Shizuoka 411-8540, Japan
          Tel: 81-559-81-6856
          Fax: 81-559-81-6855
          Email: tshini@genes.nig.ac.jp.
          Location/Qualifiers
            1..517
              /organism="Triticum aestivum"
              /cultivar="Chinese Spring"

```

```

/db_xref="taxon:4565"
/clone_lib="Y. Ogihara unpublished cDNA library, Wh_e"
/tissue_type="seed DPAL0"
/dev_stage="Feekes' scale 11.2"
/notes="Vector: Lambda Uni-ZAP XR, excised phagemid;
Site1: EcoRI; Site2: XhoI; Plants were grown under
hydroponic conditions at UC Davis, salt stressed for 12
hours, and for 7 days, then dissected and frozen (Akhunov
in J Dvorak Lab). Total RNA was prepared from sheath
tissue, equal quantities of RNA were pooled from the two
samples, polyA was purified from the pooled RNA, a cDNA
library was made, and the cDNA clones were in vivo
excised to give pBluescript phagemids in the TJ Close lab
at the University of California, Riverside (Akhunov, Chin
, Choi, Close, Fenton, Kianian, Otto, Simons, Zhang).
Plasmid DNA preparations and DNA sequencing were
performed in the OD Anderson lab (all other authors)."
BASE COUNT      134 a 138 c 158 g 85 t
ORIGIN

Query Match      22.6%; Score 47; DB 13; Length 517;
Best Local Similarity 57.0%; Pred. No. 0.93;
Matches 86; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

QY 23 GGCCCGGTAGCGGCTACAAAGGTCAACAAGAGCTCGGCTCGCGGATCCGAT 82
|||||
Db 191 GACGCGGTGGCGCGCTACGAGCTCGGCAAGACCGTCGGCGAGGAGCTTCCCAAGGT 250
|||||

QY 83 CACCAGCTCCACGCTGACCGAGGAGACGTCGTCGCCACCATCGATCTGTCTCGCCT 142
|||||
Db 251 CAAGATCCCAAGACACCCGCAACGCGCACCTCGGCCATCAAGGTGCTCGACGCAA 310
|||||

QY 143 GCACGAGGCGCAGCACGATGACCGTCCCG 173
|||||
Db 311 CCAGTCTCCGCGCAAGATGTCGAGCAG 341
|||||

RESULT 3
BG908023
LOCUS      TaLr1164F05R TaLr1 Triticum aestivum cDNA clone TaLr1164F05 5',
DEFINITION mRNA sequence.
ACCESSION  BG908023
VERSION     BG908023.1 GI:14315686
KEYWORDS   EST.
SOURCE     bread wheat.
ORGANISM   Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
; Triticeae; Triticum.
REFERENCE  1 (bases 1 to 649)
AUTHORS   Cloutier, S., Dong, G. and Walsh, A.
TITLE     Wheat functional genomics - Thatcher Lr1 cDNA library
JOURNAL   Unpublished (2001)
COMMENT   Contact: Dr. Sylvie Cloutier
          Cereal Research Centre, Agriculture and Agri-food Canada
          195 Dafoe Rd, Winnipeg, MB, Canada R3T 2N9
          Tel: (204) 983-2340
          Fax: (204) 983-4604
          Email: scloutier@em.agr.ca
          was cloned directionally, not all sequences generated with reverse
          primer were from the 5' end (same with forward primer and 3' end).
          Average inset size is >2.2 kb
          Plate: 164 row: F column: 05
          Seq primer: M13 Reverse.
          Location/Qualifiers
            1..649
              /organism="Triticum aestivum"
              /cultivar="Thatcher Lr1"
              /db_xref="taxon:4565"
              /clone="TaLr1164F05"
              /clone_lib="TaLr1"

```

```
/tissue_type="Leaf tissue"
/dev_stage="14 Days old"
/lab_host="E. coli XL0LR"
/note="Vector: Lambda ZapII; mass excised in plasmid
vector pBK-CMV (Stratagene).; Site_1: EcoRI; Site_2: XhoI;
mRNA obtained from wheat NIL Thatcher lrl 24 hours after
inoculation with leaf rust pathogen Puccinia triticina
race BBB carrying the avirulence gene Avr1."
BASE COUNT      162 a  178 c  179 g  130 t
ORIGIN

Query Match      22.6%; Score 47; DB 12; Length 649;
Best Local Similarity 57.0%; Pred. No. 0.97;
Matches 86; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

QY 23 GCGCGGGTAGGCGCTACAAGGTCACAAGAGCTGGCGCTGCACGTGCGCGATCGAT 82
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 188 GACGCGGGTGGGCCCTACGAGCTCGGCAAGACCGCTGGCGGAGGCGAGCTTCGCCAAGGT 247
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 83 CACCAGCTCCACGCTGACGAGGAGAGCTGCTGCCACCATCATGCTGCTGCCGCT 142
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 248 CAAGATCCCAAGGACACCCGACGCGCCACCTGCGCCATCAAGGTGCTCGACCCAA 307
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 143 GCACGAGGCGCACGACGATGACCGTCCCG 173
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 308 CCACGTCTCCGCCACAGATGGTCGAGCAG 338
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 4
BJ258518
LOCUS
DEFINITION      BJ258518 Y. Oghara unpublished cDNA library, wh_h Triticum
aestivum cDNA clone whh8d07 5', mRNA sequence.
ACCESSION      BJ258518
VERSION        BJ258518.1
KEYWORDS       EST.
SOURCE         bread wheat.
ORGANISM       Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae
; Triticeae; Triticum.
REFERENCE      1 (bases 1 to 684)
AUTHORS        Oghara,Y. and Murai,K.
TITLE          Expressed genes in Triticum aestivum
JOURNAL        Unpublished (2002)
COMMENT        Contact: Tadasu Shin-i
                Center For Genetic Resource Information
                National Institute of Genetics
                1111 Yata, Mishima, Shizuoka 411-8540, Japan
                Tel: 81-559-81-6856
                Fax: 81-559-81-6855
                Email: tshini@genes.nig.ac.jp.
                Location/Qualifiers
                1..684
                /organism="Triticum aestivum"
                /cultivar="Chinese Spring"
                /db_xref="taxon:4565"
                /clone="whh8d07"
                /clone_lib="Y. Oghara unpublished cDNA library, wh_h"
                /tissue_type="spike at heading date"
                /dev_stage="Feekes' scale 10.5"
                /note="Vector: Lambda Uni-ZAP XR, excised phagemid;
                Site_1: EcoRI; Site_2: XhoI; Plants were grown under
                hydroponic conditions at UC Davis, salt stressed for 12
                hours, and for 7 days, then dissected and frozen (Akhunov
                in J Dvorak Lab). Total RNA was prepared from sheath
                tissue, equal quantities of RNA were pooled from the two
                samples, polyA was purified from the pooled RNA, a cDNA
                library was made, and the cDNA clones were in vivo
                excised to give pBluescript phagemids in the TJ Close lab
                at the University of California, Riverside (Akhunov, Chin
                , Choi, Close, Fenton, Kianian, Otto, Simons, Zhang).
                Plasmid DNA preparations and DNA sequencing were
```

```
performed in the OD Anderson lab (all other authors).
BASE COUNT      180 a  178 c  190 g  136 t
ORIGIN

Query Match      22.6%; Score 47; DB 13; Length 684;
Best Local Similarity 57.0%; Pred. No. 0.97;
Matches 86; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

QY 23 GCGCGGGTAGGCGCTACAAGGTCACAAGAGCTGGCGCTGCACGTGCGCGATCCGAT 82
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 178 GACGCGGGTGGGCCCTACGAGCTCGGCAAGACCGCTGGCGGAGGCGAGCTTCGCCAAGGT 237
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 83 CACCAGCTCCACGCTGACGAGGAGAGACGCTGCTGCCACCATCATGCTGCTGCCGCT 142
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 238 CAAGATCGGCAAGGACACCCGCAACCGCCACCTGCGCCATCAAGGTGCTCGACCGCA 297
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 143 GCACGAGGCGCACGACGATGACCGTCCCG 173
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 298 CCACGTCTCCGCCACAGATGGTCGAGCAG 328
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 5
BE418320
LOCUS
DEFINITION      BE418320 1013 bp mRNA linear EST 24-JUL-2000
cDNA clone SCL024.F08 Wheat Leaf Library Triticum aestivum
ACCESSION      BE418320
VERSION        BE418320.1
KEYWORDS       GI:9416166
SOURCE         bread wheat.
ORGANISM       Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae
; Triticeae; Triticum.
REFERENCE      1 (bases 1 to 1013)
AUTHORS        Anderson,O.A., Appels,R., Bailey,P., Blake,T., Close,T., Cloutier
                ,S., Dubcovsky,J., Feuillet,C., Gale,M., Graner,A., Gustafson,P.,
                Herrmann,R.G., Holton,T., Jacquemin,J.M., Jia,J., Joudrier,P.,
                Langridge,P., Lazo,G.R., Lin,J.J., McGuire,P., Ogihara,Y.,
                Pecchioni,N., Qualset,C., Schuch,W., Selvaraj,G., Shariflou,M.,
                Sorrells,M., Warburton,M. and Wenzel,G.
                International Triticeae EST Cooperative (ITEC): Production of
                Expressed Sequence Tags for Species of the Triticeae
                Unpublished (2000)
                Contact: Cloutier S
                Cereal Research Centre, Agriculture & Agri-Foods Canada
                Winnipeg MT CANADA
                Tel: 204 983 2340
                Fax: 204 983 4604
                Email: scloutier@em.agr.ca
                International Triticeae EST Cooperative (ITEC)
                http://wheat.pw.usda.gov/genome.
                Location/Qualifiers
                1..1013
                /organism="Triticum aestivum"
                /cultivar="Thatcher Lrl"
                /db_xref="taxon:4565"
                /clone="SCL024.F08"
                /clone_lib="ITEC SCL Wheat Leaf Library"
                /tissue_type="young leaf"
                /dev_stage="after 24 hour challenge with LR-AVR1"
                /note="Vector: Lambda ZAP; 1.0 Kbp average insert size."
BASE COUNT      246 a  237 c  271 g  223 t  36 others
ORIGIN

Query Match      22.6%; Score 47; DB 10; Length 1013;
Best Local Similarity 57.0%; Pred. No. 1;
Matches 86; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

QY 23 GCGCGGGTAGGCGCTACAAGGTCACAAGAGCTGGCGCTGCACGTGCGCGATCCGAT 82
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 184 GACGCGGGTGGGCCCTACGAGCTCGGCAAGACCGCTGGCGGAGGCGAGCTTCGCCAAGGT 243
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```











TITLE A BAC End Sequencing Framework to Sequence the Rice Genome  
JOURNAL Unpublished (1998)  
COMMENT Contact: Wing RA  
Clemson University Genomics Institute  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA  
Tel: 864 656 7288  
Fax: 864 656 4293  
Email: rwing@clemson.edu  
Seq primer: TAATACGACTCATATAGGG  
Class: BAC ends  
High quality sequence stop: 85.  
Location/Qualifiers  
source  
1. .877  
/organism="Oryza sativa"  
/strain="Japonica"  
/cultivar="Nipponbare"  
/db\_xref="taxon:4530"  
/clone="nxb0075L01f"  
/clone\_lib="CUGI Rice BAC Library"  
/tissue\_type="Leaf"  
/lab\_host="E. coli DH10B"  
/note="Vector: pBelobAC11; Site\_1: HindIII; Site\_2:  
HindIII; Rice is one of two most popular grains in the  
world. Half of the world population especially those  
inhabiting highly populated areas of the humid tropics  
and subtropics, rely on rice as their primary source of  
carbohydrate. Monocotyledonous rice is a diploid plant  
(2n=24) with a haploid genome equivalent of 431 Mbp  
(Arumuganathan and Earle, 1991). The relatively small  
genome of rice, three times larger than that of  
Arabidopsis, makes it suitable for genomic studies. In  
order to facilitate positional cloning, physical mapping  
and genome sequencing of rice, we have constructed a BAC  
library from Oryza sativa, Nipponbare variety. The  
library contains 36,864 clones with an average insert size  
of 128.5 kb providing 10.9 haploid genome equivalents. The  
deep coverage allows the isolation a particular sequence  
with a probability of 99.9 %. Two high density filters,  
each containing 18,432 clones (doubly spotted), represent  
the whole library for colony screening."  
BASE COUNT 107 a 370 c 257 g 143 t  
ORIGIN  
Query Match 21.3%; Score 44.2; DB 17; Length 877;  
Best Local Similarity 53.8%; Pred. No. 4.1;  
Matches 91; Conservative 0; Mismatches 78; Indels 0; Gaps 0;  
Qy 33 GGCGCTACAAAGTCAACAGAGCTCGCGCTGCAGCTCGCGCATCCGATCACCAGCTCC 92  
Db 232 GCGCGGAGGCGACCAACCGAGCGCGTGTGACCGTGTGCTCGAGTTCCTCGCGACGTT 173  
Qy 93 ACCTGACCGAGGAGACGTGCTGGCCACCATCGATACCTGCTGCGCTGACAGAGGC 152  
Db 172 CAGGGATGGAGCGGAGCTGGAGCGGACATGGAGCGCATGGGCGCCCGAGACGAGCG 113  
Qy 153 CAGCACAGTACCGCTCCCGGGCGGACCGAGGTGCGCGTTCGAGACCG 201  
Db 112 CTGCTCGGCGGACGACGCGCGCTGGGGCGGAGCTGCCGAGGGG 64  
RESULT 15  
BI995346  
LOCUS BI995346 568 bp mRNA linear EST 25-OCT-2001  
DEFINITION 1031026E07.y2 C. reinhardtii CC-1690, Stress II (normalized),  
Lambda Zap II Chlamydomonas reinhardtii cDNA, mRNA sequence.  
ACCESSION BI995346  
VERSION BI995346.1 GI:16430136  
KEYWORDS EST.  
SOURCE Chlamydomonas reinhardtii.  
ORGANISM Chlamydomonas reinhardtii.  
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;  
Chlamydomonadaeae; Chlamydomonas.

REFERENCE 1 (bases 1 to 568)  
AUTHORS Grossman,A., Chang,C.-W., Davies,J., Harris,E., Hauser,C., Lefebvre  
P., McDermott,J.P., Shrager,J., Silflow,C. and Stern,D.  
TITLE Analyses of the Chlamydomonas reinhardtii Genome: A Model,  
Unicellular System for Analyzing Gene Function and Regulation in  
Vascular Plants. Project: 1031  
JOURNAL Unpublished (2001)  
COMMENT Contact: Charles Hauser  
DCMB Box 91000  
Duke University  
Durham, NC 27708-1000  
Tel: 919 613 8159  
Fax: 919 613 8177  
Email: chauser@duke.edu.  
Location/Qualifiers  
source  
1. .568  
/organism="Chlamydomonas reinhardtii"  
/strain="CC-1690 wild type mt+ 21gr"  
/db\_xref="taxon:3055"  
/clone\_lib="C. reinhardtii CC-1690, Stress II (normalized  
) , lambda Zap II"  
/note="vector: pBluescript II SK-; Site\_1: EcoRI; Site\_2:  
XhoI; Stress condition II library, constructed by John  
Davies and Jeffrey McDermott, combines cDNAs from CC-1690  
cells grown to mid-log phase in TAP (NH4+ - containing)  
and shifted to TAP - NO3- (24hrs); H2 production  
conditions (0, 12hr, 24hr) see Melis et al., (2000) Plant  
Phys. 122: 127-135; TAP + H2O2 (1, 12, 24 hr); TAP +  
sorbitol (1, 2, 6, 24 hr); TAP + Cd (1, 2, 6, 24 hr).  
POLYA mRNA was purified from each sample, pooled and cDNA  
synthesized. The cDNA was directionally cloned into lambda  
Zap II (Stratagene) in the EcoRI (5') and XhoI (3')  
sites. pBluescript II SK- plasmids were excised from the  
lambda Zap clones by superinfection with Exassist  
(Stratagene) phage. The library was normalized using  
method 4 described in Bonaldo et al., (1996) Genome  
Research 6: 791-806."  
BASE COUNT 80 a 206 c 185 g 97 t  
ORIGIN  
Query Match 21.1%; Score 43.8; DB 13; Length 568;  
Best Local Similarity 50.7%; Pred. No. 4.7;  
Matches 105; Conservative 0; Mismatches 102; Indels 0; Gaps 0;  
Qy 2 CAAGGAGAGCGCTACGACCTGGCCCGGTAGCGCGCTACAAGGTCAACAGACCTCGG 61  
Db 251 CAAGGAGAGCGGCTTCGAGGTGGTGCACGCGCTGTGAACAGGGTCAATGAGGTGCTGG 310  
Qy 62 CCTGCACGCTCGCGCATCCGATCACCAGCTCCACGCTGACCGAGGAGACGTCGTCCGCC 121  
Db 311 CGTGCCTACAGGAGAGCGCGCGCTCTCTCCGCGCCGCTGCTCTACCACTGGCGCC 370  
Qy 122 CATCGAGTACCTGGTCCGCCCTGACGAGGGGCGACACAGATGACCGTCCCGGGGGCGAC 181  
Db 371 CGCCGCCGACCCCGCGCTGTTCGCCGGCGGCGCGCGGTGTTCCGCTCCACCTCGCG 430  
Qy 182 CGAGTGGCGGTTGAGACCGAGCAT 208  
Db 431 CGCAAGAGCCCGTGGCGGCTGTTCCGGCAT 457  
Search completed: November 13, 2002, 03:59:26  
Job time : 1134.63 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 12, 2002, 16:05:40 ; Search time 21.1607 Seconds  
(without alignments)  
3487.380 Million cell updates/sec

Title: US-09-697-123B-1

Perfect score: 208

Sequence: 1 tcaagagaagcgtacgac.....coggttgagaccgacacat 208

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 320260 seqs, 177392727 residues

Total number of hits satisfying chosen parameters: 640520

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Published Applications\_NA:\*
- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*
  - 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*
  - 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*
  - 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*
  - 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*
  - 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*
  - 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*
  - 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*
  - 9: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*
  - 10: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*
  - 11: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*
  - 12: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*
  - 13: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*
  - 14: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query No.	Score	Match	Length	DB ID	Description
1	83.2	40.0	5096	10	US-09-984-711-5	Sequence 5, Appli
2	83.2	40.0	5099	9	US-10-075-460-5	Sequence 5, Appli
3	83.2	40.0	5099	10	US-09-887-052-1	Sequence 1, Appli
4	83.2	40.0	5099	10	US-09-887-052-3	Sequence 3, Appli
5	83.2	40.0	5099	10	US-09-887-052-5	Sequence 5, Appli
6	44.4	21.3	2329	10	US-09-816-828-9	Sequence 9, Appli
7	38.4	18.5	1998	10	US-09-815-242-7695	Sequence 7695, Ap
8	37.2	17.9	1066	10	US-09-879-833-1	Sequence 1, Appli
9	37.2	17.9	1437	10	US-09-815-242-7738	Sequence 7738, Ap
10	37.2	17.9	2451	10	US-09-771-161A-4	Sequence 4, Appli
11	37.2	17.9	4444	10	US-09-879-833-3	Sequence 3, Appli
12	36.6	17.6	526	10	US-09-731-872-190	Sequence 190, App
13	36.6	17.6	654	10	US-09-731-872-229	Sequence 229, App
14	36.4	17.5	1314	10	US-09-815-242-7740	Sequence 7740, Ap
15	36.2	17.4	427	10	US-09-960-352-5073	Sequence 5073, Ap
16	36.2	17.4	978	10	US-09-815-242-8005	Sequence 8005, Ap
17	36.2	17.4	1266	10	US-09-815-242-7920	Sequence 7920, Ap
18	36.2	17.4	3731	10	US-09-832-268A-1	Sequence 1, Appli
19	36	17.3	1004	10	US-09-755-830-7	Sequence 7, Appli

ALIGNMENTS

RESULT 1

US-09-984-711-5  
; Sequence 5, Application US/09984711  
; Patent No. US20020119549A1

GENERAL INFORMATION:

; APPLICANT: MOECKEL, Bettina  
; APPLICANT: BATHE, Brigitte  
; APPLICANT: STEPHAN, Hans  
; APPLICANT: KREUTZER, Caroline  
; APPLICANT: HERMANN, Thomas  
; APPLICANT: PFEFFERLE, Walter  
; APPLICANT: BINDER, Michael  
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES WHICH CODE FOR THE rpsL GENE  
; FILE REFERENCE: 204209US0  
; CURRENT APPLICATION NUMBER: US/09/984,711  
; CURRENT FILING DATE: 2001-10-31  
; PRIOR APPLICATION NUMBER: DE10108230.9  
; PRIOR FILING DATE: 2001-02-16  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 5  
; LENGTH: 5096

; TYPE: DNA  
; ORGANISM: Corynebacterium glutamicum

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (702)..(4196)

OTHER INFORMATION:

US-09-984-711-5

Query Match 40.0% Score 83.2; DB 10; Length 5096;

Best Local Similarity 67.0%; Pred.No. 1.8e-12;

Matches 134; Conservativ: 0; Mismatches 63; Indels 3; Gaps 1;

Qy 9 AAGCGCTACGACCTGGCCGGGTAGCGGTACAAAGTCAACAAAGCTGGCGCTGCAC 68

|||||

Db 1578 AAGCGCTACGACCTGGCGGTGGTGGTTCACAAAGTCAACCGCAAGCTCGGCCTT--- 1634

|||||

Qy 69 GTCGGCGATCGATCACCAGCTCCACGCTGACCGAGGAGGAGCTGCGCGCCACCATCGAG 128

|||||

Db 1635 GGTGGCGACCATGATGTTGTTGATGACTCTTACTGAAGGAGGACATCGCAACCATCGAG 1694

|||||







SOFTWARE: PatentIn version 3.0

SEQ ID NO 4  
LENGTH: 2451  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: -  
LOCATION: (1)..(2451)  
OTHER INFORMATION: "n" can be any nucleotide 'a', 'c', 'g' or 't'  
US-09-771-161A-4

Query Match 17.9%; Score 37.2; DB 10; Length 2451;  
Best Local Similarity 53.4%; Pred. No. 0.4;  
Matches 78; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 28 GGGTAGCCGCTACAGGCTCAACAAGAGCTCGCGCTGCACGTCGGCGATCCGATCACCA 87  
DB 1748 GGATGAGCAACCCCGAGGTCTCCGCAACCTGGAGCGGGTACCGCATGCGCGGCCCG 1807  
QY 88 GTCCACGCTGACCCGAGGAAGAGCTGCTGCCACCACATCGATACCTGGTCCGCGCTGCACG 147  
DB 1808 ACACCTGCCCGCCGAGCTGTACCGCGGGTCTATCGCGGAGTGTGGCGCAGCCGCCG 1867  
QY 148 AGGGCAGCACAGCATGACCGTCCCG 173  
DB 1868 AGGAGCGGCCACCTTCGAGTTCCCTG 1893

RESULT 11

US-09-879-833-3  
Sequence 3, Application US/09879833  
Patent No. US2002055107A1

GENERAL INFORMATION:  
APPLICANT: Park Chong-Hun; SAMYANG GENEX CORPORATION  
TITLE OF INVENTION: CANCER DIAGNOSIS METHOD USING CELL GROWTH INHIBITING AND CELL DIFFERENTIATION SPECIFIC SYG972 GENE AND GENOMIC DNA AND PROMOTER T  
FILE OF INVENTION: FERENTATION  
FILE REFERENCE: PA/SYG/99601  
CURRENT FILING DATE: 2001-06-12  
PRIOR APPLICATION NUMBER: KR 54933  
PRIOR FILING DATE: 1998-12-12  
PRIOR APPLICATION NUMBER: KR 63958  
PRIOR FILING DATE: 1998-12-30  
NUMBER OF SEQ ID NOS: 3  
SOFTWARE: KOPATIN 1.5  
SEQ ID NO 3  
TYPE: DNA  
ORGANISM: HL60 cell line  
US-09-879-833-3

Query Match 17.9%; Score 37.2; DB 10; Length 4444;  
Best Local Similarity 52.6%; Pred. No. 0.42;  
Matches 81; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 55 AGCTGGCCTGCACGTCGGCGATCCGATCACCGCTCCACGCTGACCGAGGAAGACGTCG 114  
DB 3669 AGAGGACAGCGGCGACATCGCGCTGCAGATCCATTTTACGCTGATCCAGGCTTTCTGCT 3728  
QY 115 TCGCCACCATCGATACCTGTGTCGCGCTGCACGAGGGCCGACGACGATGACCGTCCCGG 174  
DB 3729 GCGAGAACGACATGACATAGTGCCTGGCGGTGATGTGCGGCGGTGCGGCTATCGTGG 3788  
QY 175 GCGGACCGAGTGGCGGTGTAGACCGGAGGACAT 208  
DB 3789 GCGCGCGGAGGAGCGGCTGCGCGGCGACCT 3822

RESULT 12

US-09-731-872-190  
Sequence 190, Application US/09731872  
Patent No. US20020102604A1

GENERAL INFORMATION:

APPLICANT: Dumas Milne Edwards, Jean Baptiste  
APPLICANT: Bougueleret, Lydie  
APPLICANT: Jobert, Severin  
TITLE OF INVENTION: FULL-LENGTH HUMAN CDNAS ENCODING POTENTIALLY SECRETED PROTEINS  
FILE REFERENCE: 78.US3.REG  
CURRENT APPLICATION NUMBER: US/09/731,872  
CURRENT FILING DATE: 2000-12-07  
PRIOR APPLICATION NUMBER: US 60/169,629  
PRIOR FILING DATE: 1999-12-08  
PRIOR APPLICATION NUMBER: US 60/187,470  
PRIOR FILING DATE: 2000-03-06  
NUMBER OF SEQ ID NOS: 482  
SOFTWARE: Patent.pm  
SEQ ID NO 190  
LENGTH: 526  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 3..368  
US-09-731-872-190

Query Match 17.6%; Score 36.6; DB 10; Length 526;  
Best Local Similarity 54.0%; Pred. No. 0.51;  
Matches 75; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 22 TGGCCCGGGTAGCGCTACAGGCTCAACAAGAGCTCGCGCTGCACGTCGGCGATCCGA 81  
DB 64 TGGCTGGGGCGCGCTACCTGCTGTACGACCAAGAGCTGCTGGGCGCCAGGACAGA 123  
QY 82 TCACCACTCCACGCTCAGCAGGAGAACACCTGCTGCCACCATCGACTACTGTCGCC 141  
DB 124 GCCAGGAGCCCTACAGAGCTGGGGAGGTGGTCCCCCGCCCATGTACCACTGACGC 183  
QY 142 TGCACGAGGCGCCAGCACAC 160  
DB 184 AGTACGTGTGTACGACAGAC 202

RESULT 13

US-09-731-872-229  
Sequence 229, Application US/09731872  
Patent No. US20020102604A1

GENERAL INFORMATION:  
APPLICANT: Dumas Milne Edwards, Jean Baptiste  
APPLICANT: Bougueleret, Lydie  
APPLICANT: Jobert, Severin  
TITLE OF INVENTION: FULL-LENGTH HUMAN CDNAS ENCODING POTENTIALLY SECRETED PROTEINS  
FILE REFERENCE: 78.US3.REG  
CURRENT APPLICATION NUMBER: US/09/731,872  
CURRENT FILING DATE: 2000-12-07  
PRIOR APPLICATION NUMBER: US 60/169,629  
PRIOR FILING DATE: 1999-12-08  
PRIOR APPLICATION NUMBER: US 60/187,470  
PRIOR FILING DATE: 2000-03-06  
NUMBER OF SEQ ID NOS: 482  
SOFTWARE: Patent.pm  
SEQ ID NO 229  
LENGTH: 654  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 86..505  
US-09-731-872-229

Query Match 17.6%; Score 36.6; DB 10; Length 654;  
Best Local Similarity 54.0%; Pred. No. 0.52;  
Matches 75; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 22 TGGCCCGGGTAGCGCTACAGGCTCAACAAGAGCTCGCGCTGCACGTCGGCGATCCGA 81  
DB 184 AGTACGTGTGTACGACAGAC 202



Db 201 TGGCTGGGGCCCGCTTACCTGGTGTACGACCGAGGAGCTGCTGGGGCCCGCAGACAAGA 260  
QY 82 TCACGAGCTCCACGCTGACCGAGGAAGAGCTGCTGCCGACCATCGAGTACCTGGTCCGCC 141  
Db 261 GCCAGGACCCCTACAGAGGCTGGGAGGTGGTCCCGCCGCCATGTACCAGTTCAGCC 320  
QY 142 TCACGAGCGCCGAGCAGAC 160  
Db 321 AGTACGTGTGTCAGCAGAC 339

RESULT 14

US-09-815-242-7740  
; Sequence 7740, Application US/09815242  
; Patent No. US20020061569A1  
; GENERAL INFORMATION:  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Karl L.  
; APPLICANT: Zyskind, Judith W.  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John D.  
; APPLICANT: Carr, Grant J.  
; APPLICANT: Yamamoto, Robert T.  
; APPLICANT: Xu, H. Howard  
; TITLE OF INVENTION: Identification of Essential Genes in  
; TITLE OF INVENTION: Prokaryotes  
; FILE REFERENCE: ELITRA.011A  
; CURRENT APPLICATION NUMBER: US/09/815,242  
; CURRENT FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; NUMBER OF SEQ ID NOS: 14110  
; SOFTWARE: FastSEQ for Windows version 4.0  
; SEQ ID NO 7740  
; LENGTH: 1314  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)...(1314)  
US-09-815-242-7740

Query Match 17.5%; Score 36.4; DB 10; Length 1314;  
Best Local Similarity 51.9%; Pred. No. 0.61;  
Matches 82; Conservative 0; Mismatches 76; Indels 0; Gaps 0;  
QY 26 CCGGGTAGCGCGCTACAGGTCAACAGAGCTGCGGCTGCGAGTCCGGATCCGATCAC 85  
Db 612 CGGTATCGTATCGCCACCGCCATTGCGCACCGCGGCATCAGCGACGCGGCTGGTGCAC 671  
QY 86 CAGCTCCACCGCTACCGAGGAAGAGCTGCTGCCGACCATCGAGTACCTGGTCCGCTGCA 145  
Db 672 GCGCTATTCATGATCGCCGGGGTTCATCGCCGCCACCGCGGCTGTCTTACCTGGC 731  
QY 146 CGAGGGCCAGCACAGCATGATACCGTCCCGGGCGGCACCG 183  
Db 732 GCTGTTCTATCTTGGCGCCACAGCCAGGCGCATCGCG 769

RESULT 15

US-09-960-352-5073

; Sequence 5073, Application US/09960352  
; Patent No. US20020137139A1  
; GENERAL INFORMATION:  
; APPLICANT: Warren, Wesley C.  
; APPLICANT: Tao, Nengbing  
; APPLICANT: Byatt, John C.  
; APPLICANT: Mathialagan, Nagappan  
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND  
; FILE REFERENCE: 16511.006/37-21(10298)C  
; CURRENT APPLICATION NUMBER: US/09/960,352  
; CURRENT FILING DATE: 2001-09-24  
; NUMBER OF SEQ ID NOS: 15112  
; SEQ ID NO 5073  
; LENGTH: 427  
; TYPE: DNA  
; ORGANISM: Bos taurus  
; OTHER INFORMATION: Clone ID: 22-LIB3058-020-Q1-K1-F9  
US-09-960-352-5073

Query Match 17.4%; Score 36.2; DB 10; Length 427;  
Best Local Similarity 49.2%; Pred. No. 0.63;  
Matches 95; Conservative 0; Mismatches 98; Indels 0; Gaps 0;  
QY 8 GAAGCGCTAGACCTGGCCCGGTAGCGCTACAGGTCAACAGAGCTCGCCTGCA 67  
Db 26 GCACAGCGGGGCTTCAACCTGGAGGCTGCCGCGAGTACAGCCGCTCATGCTTACAA 85  
QY 68 CGTCGGGATCCGATCACCAGCTCCAGCTGACCGAGGAAGAGCTGTCGCCACCATCGA 127  
Db 86 CGCATCGACTCGCTGACGCGCATCGTGGCGGCCCTGGCGCGCTCAAGATCGACTTCCA 145  
QY 128 GTACCTGTGCGCTGACGAGGCGCCAGCACAGCATGACCGTCCCGGGCGGCACCGAGGT 187  
Db 146 CAACCCCGACCGCGCTACGACCGCGCTGTCGCGCTACGCGGCGCCGCGAGAG 205  
QY 188 GCCGTTGAGACC 200  
Db 206 CAAGGGCGAGATC 218

Search completed: November 12, 2002, 16:57:36  
Job time : 35.1607 secs

Result No.	Query			ID	Description		
	Score	Match	Length				
C	1	166.4	80.0	970	1	US-08-250-030-1	Sequence 1, Appli
	2	166.4	80.0	970	5	PCF-US95-06790-1	Sequence 1, Appli
	3	142.4	68.5	3447	2	US-08-313-185-57	Sequence 57, Appli
	4	142.4	68.5	3447	3	US-09-082-6114A-57	Sequence 57, Appli
	5	49.8	23.9	6085	4	US-09-029-603-4	Sequence 4, Appli
	6	45.4	21.8	4403765	4	US-09-103-840A-2	Sequence 2, Appli
	7	43	20.7	2122	4	US-09-029-603-1	Sequence 1, Appli
	8	42	20.2	2244	1	US-08-476-519-10	Sequence 10, Appli
	9	42	20.2	2244	5	PCF-US95-09323-10	Sequence 10, Appli
	10	42	20.2	2334	1	US-08-476-519-1	Sequence 1, Appli
C	11	42	20.2	2334	5	PCF-US95-09323-1	Sequence 1, Appli
	12	40.2	19.3	4403765	4	US-09-103-840A-2	Sequence 2, Appli
	13	40.2	19.3	4411529	4	US-09-103-840A-1	Sequence 1, Appli
	14	39.2	18.8	12588	2	US-08-387-942C-1	Sequence 1, Appli
	15	39	18.8	14377	2	US-08-804-227C-7	Sequence 7, Appli
	16	39	18.8	44377	2	US-08-804-198-1	Sequence 1, Appli
	17	38.4	18.5	1998	4	US-09-382-106-1	Sequence 1, Appli
	18	38	18.3	765	3	US-07-718-904-79	Sequence 79, Appli
	19	37.4	18.0	20235	1	US-07-642-734C-3	Sequence 3, Appli
	20	37.4	18.0	20235	3	US-08-439-009A-3	Sequence 3, Appli
C	21	37.2	17.9	477	1	US-08-797-831A-2	Sequence 2, Appli
	22	37.2	17.9	477	2	US-08-463-081B-31	Sequence 31, Appli
	23	37.2	17.9	477	2	US-08-461-379A-31	Sequence 31, Appli
	24	37.2	17.9	477	2	US-08-462-390B-31	Sequence 31, Appli
	25	37.2	17.9	477	3	US-08-463-074B-31	Sequence 31, Appli
	26	37.2	17.9	477	3	US-08-465-585C-31	Sequence 31, Appli
	27	37.2	17.9	477	3	US-08-652-446-31	Sequence 31, Appli

```
QY 61 GCCTGCACGTCGGGATCCGATCCACAGCTCCACGCTGACCGAGGAGAGCTGCTCGGCCA 120
Db 86 GGCTGCATGTCGGGAGCCCATCAGCTGCTGACGCTGACCGGAAGAGAGCTGCTGGCCA 145
QY 121 CCATCGAGTACCTGGTCCGCTGCACGAGGCCAGCACAGATGACCGTCCCGGGCGGCA 180
Db 146 CCATCGAATATCTGGTCCGCTGCACGAGGTCAGACCCAGATGACCGTTCCGGGGCGG 205
QY 181 CCGAGGTCCGGTTGACACGAGACAT 208
Db 206 TCGAGGTGCGCGTGGAACCGAGACAT 233

RESULT 2
PCT-US95-06790-1
: Sequence 1, Application PC/TUS9506790
: GENERAL INFORMATION:
: APPLICANT: Mayo Foundation for Medical Education and Research
: APPLICANT: and Hoffmann-La Roche Inc.
: TITLE OF INVENTION: Detection of a Genetic Locus Encoding
: TITLE OF INVENTION: Resistance to Rifampin
: NUMBER OF SEQUENCES: 15
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Schwegman, Lundberg & Woessner
: STREET: 3500 IDS Center
: CITY: Minneapolis
: STATE: MN
: COUNTRY: USA
: ZIP: 55402
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: PCT/US95/06790
: FILING DATE: 26-MAY-1995
: CLASSIFICATION:
: ATTORNEY/AGENT INFORMATION:
: NAME: Raasch, Kevin W.
: REGISTRATION NUMBER: 35,651
: REFERENCE/DOCKET NUMBER: 150.105W01
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 612-339-0331
: TELEFAX: 612-339-3061
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 970 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA
PCT-US95-06790-1

Query Match 80.0%; Score 166.4; DB 5; Length 970;
Best Local Similarity 87.5%; Pred. No. 4e-30;
Matches 182; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 1 TCAAGGAGAGCCCTACGACCTGGCCGGTAGGCGCTACAGGTCACACAGAGCTCG 60
Db 26 TCAAGGAGAGCCCTACGACCTGGCCGGTAGGCGCTACAGGTCACACAGAGCTCG 85
QY 61 GCCTGCACGTCGGGATCCGATCCACAGCTCCACGCTGACCGAGGAGAGCTGCTCGGCCA 120
Db 86 GGCTGCATGTCGGGAGCCCATCAGCTGCTGACGCTGACCGGAAGAGAGCTGCTGGCCA 145
QY 121 CCATCGAGTACCTGGTCCGCTGCACGAGGCCAGCACAGATGACCGTCCCGGGCGGCA 180
Db 146 CCATCGAATATCTGGTCCGCTGCACGAGGTCAGACCCAGATGACCGTTCCGGGGCGG 205
QY 181 CCGAGGTCCGGTTGACACGAGACAT 208
Db 206 TCGAGGTGCGCGTGGAACCGAGACAT 233
```

```
RESULT 3
US-08-313-185-57
: Sequence 57, Application US/08313185
: Patent No. 5851763
: GENERAL INFORMATION:
: APPLICANT: Heym, Beate
: APPLICANT: Cole, Stewart
: APPLICANT: Young, Douglas
: APPLICANT: Zhang, Ying
: APPLICANT: Honore, Nadine
: APPLICANT: Telenti, Amelio
: APPLICANT: Bodmer, Thomas
: TITLE OF INVENTION: Rapid Detection of Antibiotic Resistance
: TITLE OF INVENTION: in Mycobacterium Tuberculosis
: NUMBER OF SEQUENCES: 66
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
: ADDRESSEE: Dunner
: STREET: 1300 I Street, N.W.
: CITY: Washington
: STATE: D.C.
: COUNTRY: USA
: ZIP: 20005-3315
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/313,185
: FILING DATE: 12-OCT-1994
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Meyers, Kenneth J.
: REGISTRATION NUMBER: 25,146
: REFERENCE/DOCKET NUMBER: 02356.0068-00000
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (202) 408-4000
: TELEFAX: (202) 408-4400
: INFORMATION FOR SEQ ID NO: 57:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3447 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
US-08-313-185-57

Query Match 68.5%; Score 142.4; DB 2; Length 3447;
Best Local Similarity 80.3%; Pred. No. 1.4e-24;
Matches 167; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 1 TCAAGGAGAGCCCTACGACCTGGCCGGTAGGCGCTACAGGTCACACAGAGCTCG 60
Db 809 TCAAGGAGAGAGCTGACCTGGCCGGTAGGCGCTACAGGTCACACAGAGCTCG 868
QY 61 GCCTGCACGTCGGGATCCGATCCACAGCTCCACGCTGACCGAGGAGAGCTGCTCGGCCA 120
Db 869 GGTTCACGCCGGTTCAGTTGATCAGCTGCTCCAGCTGACCGAAGAGATGCTCGGCCA 928
QY 121 CCATCGAGTACCTGGTCCGCTGCACGAGGCCAGCACAGATGACCGTCCCGGGCGGCA 180
Db 929 CCATAGGATACCTGCTGCTGATGAGGTCAGTCGACAAATGACTGTCCCGAGTGGG 988
QY 181 CCGAGGTCCCGTTGAGACCCAGACAT 208
Db 989 TAGAGTCCCGAGTGAACATGACGATAT 1016

RESULT 4
US-09-082-614A-57
```

; Sequence 57, Application US/09082614A  
; Patent No. 6124098  
; GENERAL INFORMATION:  
; APPLICANT: Heym, Beate  
; APPLICANT: Cole, Stewart  
; APPLICANT: Young, Douglas  
; APPLICANT: Zhang, Ying  
; APPLICANT: Honore, Nadine  
; APPLICANT: Telenti, Amalio  
; APPLICANT: Bodmer, Thomas  
; TITLE OF INVENTION: Rapid Detection of Antibiotic Resistance  
; TITLE OF INVENTION: in Mycobacterium Tuberculosis  
; NUMBER OF SEQUENCES: 66  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &  
; ADDRESSEE: Dunner  
; STREET: 1300 I Street, N.W.  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20005-3315  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/082.614A  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/313,185  
; FILING DATE: 12-OCT-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Meyers, Kenneth J.  
; REGISTRATION NUMBER: 25,146  
; REFERENCE/DOCKET NUMBER: 02356.0068-00000  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 408-4000  
; TELEFAX: (202) 408-4400  
; INFORMATION FOR SEQ ID NO: 57:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3447 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
US-09-082-614A-57

Query Match 68.5%; Score 142.4; DB 3; Length 3447;  
Best Local Similarity 80.3%; Pred. No. 1.4e-24;  
Matches 167; Conservative 0; Mismatches 41; Indels 0; Gaps 0;  
QY 1 TCAAGGAGAGCGGTACGACCTGGCCCGGTAGGCGGTACAAAGTCAACAAAGAGCTCG 60  
Db 809 TCAAGGAGAGCGGTACGACCTGGCCCGGTAGGCGGTACAAAGTCAACAAAGAGCTCG 868  
QY 61 GCGTCAGTCGGGATCCGATCCACCTCCACGCTGACCGAGGAGACGTCGTCGCCA 120  
Db 869 GGTTCACCGCGGTGAGTGTGATCGTCCACGCTGACCGAGAGAGATGTCGTCGCCA 928  
QY 121 CCATCGAGTACCTGGTCCGCTCGACGAGGCGCCAGCACGATGACCGCTCCCGGCGGCA 180  
Db 929 CCATAGAGTACCTGGTTCGTGTCATGAGGTCAGTCGACATGACTGTCGCCAGTGGG 988  
QY 181 CCGAGGTCCCGTGTGAGACCGACGACAT 208  
Db 989 TAGAGTCCAGTGGAAACTGACGATAT 1016

RESULT 5  
US-09-029-603-4  
; Sequence 4, Application US/09029603

; Patent No. 6210935  
; GENERAL INFORMATION:  
; APPLICANT: Schupp, Thomas  
; APPLICANT: Engel, Natalie  
; APPLICANT: Bietenhader, Jurg  
; APPLICANT: Toupet, Christine  
; APPLICANT: Pospiech, Andreas  
; TITLE OF INVENTION: Staurosporin Biosynthesis Gene Clusters  
; FILE REFERENCE: 4-20555/A/PCT  
; CURRENT APPLICATION NUMBER: US/09/029,603  
; CURRENT FILING DATE: 1998-03-20  
; EARLIER APPLICATION NUMBER: PCT/EP96/03643  
; EARLIER FILING DATE: 1996-08-19  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 4  
; LENGTH: 6085  
; TYPE: DNA  
; ORGANISM: Streptomyces longisporoflavus  
; FEATURE:  
; NAME/KEY: misc\_RNA  
; LOCATION: (378)..(1665)  
; OTHER INFORMATION: ORF  
; FEATURE:  
; NAME/KEY: misc\_RNA  
; LOCATION: (1747)..(2553)  
; OTHER INFORMATION: ORF  
; FEATURE:  
; NAME/KEY: misc\_RNA  
; LOCATION: (2593)..(4011)  
; OTHER INFORMATION: ORF  
; FEATURE:  
; NAME/KEY: misc\_RNA  
; LOCATION: (4013)..(4999)  
; OTHER INFORMATION: ORF  
; FEATURE:  
; NAME/KEY: misc\_RNA  
; LOCATION: (5071)..(6085)  
; OTHER INFORMATION: ORF  
US-09-029-603-4

Query Match 23.9%; Score 49.8; DB 4; Length 6085;  
Best Local Similarity 60.7%; Pred. No. 0.0022;  
Matches 99; Conservative 0; Mismatches 62; Indels 2; Gaps 1;  
QY 18 GACCTGSCCGCGGTAGCGCTACAAAGTCAACAAGAGTCCGCTCGCACGTCGCGCAT 77  
Db 1017 GAGGTGACCG 1076  
QY 78 CCATCAGCAGCTCCACGCTGACCGAGGAGAGAGCGTCGTCGCCACCATGAGTACCTGGT- 136  
Db 1077 CGGACACCGGATCAGCGCTCAGCGTGGAGCGCATCGTCGCGACCTGCGTCCATCTGCTC 1136  
QY 137 -CGCCTTGCACGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 178  
Db 1137 ACCG 1179  
RESULT 6  
US-09-103-840A-2/c  
; Sequence 2, Application US/09103840A  
; Patent No. 6294328  
; GENERAL INFORMATION:  
; APPLICANT: FLEISCHMAN, Robert D.  
; APPLICANT: WHITE, Owen R.  
; APPLICANT: FRASER, Claire M.  
; APPLICANT: VENTER, John C.  
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM  
; FILE REFERENCE: 24366-20007.00  
; CURRENT APPLICATION NUMBER: US/09/103,840A  
; CURRENT FILING DATE: 1998-06-24  
; NUMBER OF SEQ ID NOS: 2



```

; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Novel Isoamylase Gene, Compositions
; TITLE OF INVENTION: Containing It and Methods of Using Isoamylases
; NUMBER OF SEQUENCES: 11
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/09323
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/281902
; FILING DATE: 28-JUL-1994
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2244 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..2241
PCT-US95-09323-10

Query Match 20.2%; Score 42; DB 5; Length 2244;
Best Local Similarity 52.9%; Pred. No. 0.12;
Matches 90; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY 17 CGACCTGCCCCGGTAGCCCTACAGGTCAACAAGAGCTCGGCTGACGTGCGCGA 76
Db 1338 CGAGTGAACGGCTCTACCGCGACGGCTGCGCAAGAGCAGCAAGCTCGGCGTGA 1397
QY 77 TCCGATCACAGCTCCAGCTGACCGAGGAGACGTCGTGCCACCATCGAGTACCTGGT 136
Db 1398 AACGCTACCCCGGACCCCTGGCCAGCGCTTCGCGGGCTCCCAACGACCTGTACGCGGA 1457
QY 137 CGCGCTGCACGAGGCCAGCAGCATGACGTCCTCCGGGGCGGACCGAGG 186
Db 1458 CGAGCGCGCAAGCGGTGGCATTCGATCACTTCGTGTGTCGCCACGACG 1507

RESULT 10
US-08-476-519-1
; Sequence 1, Application US/08476519
; Patent No. 5750876
; GENERAL INFORMATION:
; APPLICANT: Barry, Gerard F.
; APPLICANT: Kishore, Ganesh M.
; TITLE OF INVENTION: No. 5750876el Isoamylase Gene, Compositions
; TITLE OF INVENTION: Containing It and Methods of Using Isoamylases
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Grace L. Bonner, Monsanto Company, BB4F
; STREET: 700 Chesterfield Parkway No. 5750876th
; CITY: St. Louis
; STATE: Missouri
; COUNTRY: USA
; ZIP: 63198
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476,519
; FILING DATE:
; CLASSIFICATION: 800

```

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/281902
; FILING DATE: 28-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Bonner, Grace L.
; REGISTRATION NUMBER: 32,963
; REFERENCE/DOCKET NUMBER: 38-21(13577)A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (314)537-7286
; TELEFAX: (314)537-6047
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2334 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-476-519-1

Query Match 20.2%; Score 42; DB 1; Length 2334;
Best Local Similarity 52.9%; Pred. No. 0.13;
Matches 90; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY 17 CGACCTGGCCCGGGTAGCGCTACAAAGGTCAACAAGAGCTCGGCTGCAGCTCGCGGA 76
Db 1431 CGAGTGAACGGCTCTACCGCGACGGCTGCGCAAGAGCAGCAAGCTCGGCGTGA 1490
QY 77 TCGATCACAGCTCCAGCTGACCGAGGAGACGTCGTGCCACCATCGAGTACCTGGT 136
Db 1491 AAGCGTCACCCCGGACCCCTGCGCACGCGCTTCGCGGGCTCCAAGACCTGTAGCGGA 1550
QY 137 CGCGCTGCACGAGGGCCAGCAGCATGACCGTCCTCCGGGGCGGACCGAGG 186
Db 1551 CGACGGCGCAAGCGGTGGCATTCGATCAACTTCGTGGTGGTCCCGCAGG 1600

RESULT 11
PCT-US95-09323-1
; Sequence 1, Application PC/TUS9509323
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Novel Isoamylase Gene, Compositions
; TITLE OF INVENTION: Containing It and Methods of Using Isoamylases
; NUMBER OF SEQUENCES: 11
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/09323
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/281902
; FILING DATE: 28-JUL-1994
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2334 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
PCT-US95-09323-1

Query Match 20.2%; Score 42; DB 5; Length 2334;
Best Local Similarity 52.9%; Pred. No. 0.13;
Matches 90; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY 17 CGACCTGGCCCGGGTAGCGCTACAAAGGTCAACAAGAGCTCGGCTGCAGCTCGCGGA 76
Db 1431 CGAGTGAACGGCTCTACCGCGACGGCTGCGCAAGAGCAGCAAGCTCGGCGTGA 1490
QY 137 CGCGCTGCACGAGGGCCAGCAGCATGACCGTCCTCCGGGGCGGACCGAGG 186
Db 1431 CGAGTGAACGGCTCTACCGCGACGGCTGCGCAAGAGCAGCAAGCTCGGCGTGA 1490

```

```
QY 77 TCCGATCACCAGCTCCACGCTGACCGAGGAAAGACGTCGTGCCACCATCATCGAGTACCTGGT 136
      ||||| | ||| ||| | | ||| | | ||| | | ||| | | ||| | | ||| | | ||| |
Db 1491 AACGGTACCCCGCCGACCCCTGGCCACCGGCTTCGCCCGCTCCACAGACCTGTAGCGCGA 1550
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 137 CCGCTGCACGAGGCCACGACACGATACCGTCCCGGGCGGCACCCGAGG 186
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1551 CGACGGCCGCAAGCGGTGCATTCGATCAACTTCGTGTGTCGCCACGACG 1600
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 12
US-09-103-840A-2
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103.840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Query Match 19.3%; Score 40.2; DB 4; Length 4403765;
Best Local Similarity 51.4%; Pred. No. 0.55;
Matches 93; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 21 CTGCGCCGGGTAGCGGCTACAAAGTCAACAAAGAGTCGGCTGCACGTCGGCGATCCG 80
      ||||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 175075 CTGCGCGGGGAGGGGCCGACGCTGTGGTCAACGACCTCGTGGCGCCCGCGGCGACG 175134
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 81 ATCACCAGCTCCACGCTGACCGAGGAGGTCGTCGCCACCATCATCGTACCTGTCGCG 140
      ||||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 175135 GCGCGCGGTTCGGCGATGCGCGATGAGGTGTCGCGCGAGATTCGCGCAAGGGGGCGG 175194
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 141 CTGCACGAGGCCACGACGATACCGTCCCGGGCGGCACCGAGGTGCCGTTGAGACC 200
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 175195 GCGGTGCGCAACTACGACAGCGTCGCCACCGAGGACGCGCAGCAACATCATCAAGACC 175254
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 201 G 201
Db 175255 G 175255

RESULT 13
US-09-103-840A-1
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103.840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
```

```
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

Query Match 19.3%; Score 40.2; DB 4; Length 4411529;
Best Local Similarity 51.4%; Pred. No. 0.55;
Matches 93; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 21 CTGCGCCGGGTAGCGGCTACAAAGTCAACAAAGAGTCGGCTGCACGTCGGCGATCCG 80
      ||||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 174906 CTGCGCGGGGAGGGGCCACGCTCGTGGTCAACGACCTCGGTGGCGCCCGCGACGCGCAG 174965
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 81 ATCACCAGTCCACGCTGACCGAGGAGAGGTGTCGCCACCATCATCGTACCTGTTGTCGCG 140
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 174966 GCGCGCGGTTCGGCGATGCGCGATGAGGTGTCGCCGAGATTGCGCAAGGGGGCGCG 175025
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 141 CTGCACGAGGCCACGACGATGACCGTCCCGGGCGGCACCGAGGTGCCGTTGAGACC 200
      ||||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 175026 GCGGTGCGCAACTACGACAGCGTCGCCACCGAGGACGCGCGCAACATCATCAAGACC 175085
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 201 G 201
Db 175086 G 175086

RESULT 14
US-08-387-942C-1
; Sequence 1, Application US/08387942C
; Patent No. 5939289
; GENERAL INFORMATION:
; APPLICANT: ERTESVAG, HELGA
; APPLICANT: VALLA, SVEIN
; APPLICANT: SKJAK-BRAEK, GUDMUND
; APPLICANT: LARSEN, BJORN
; TITLE OF INVENTION: DNA COMPOUNDS COMPRISING SEQUENCES
; TITLE OF INVENTION: ENCODING MANNURONAN C-5-EPIMERASE
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH, LLP
; STREET: P.O. BOX 747
; CITY: FALLS CHURCH
; STATE: VA
; COUNTRY: USA
; ZIP: 22042
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/387,942C
; FILING DATE: 09-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MURPHY JR, GERALD M.
; REGISTRATION NUMBER: 28,977
; REFERENCE/DOCKET NUMBER: 1809-106P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-205-8000
; TELEFAX: 703-205-8050
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12588 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Azotobacter vinelandii
; STRAIN: E
; FEATURE:
```





GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 12, 2002, 15:44:35 ; Search time 136.222 Seconds  
(without alignments)  
3438.621 Million cell updates/sec

Title: US-09-697-123B-1  
Perfect score: 208  
Sequence: 1 tcaaggagaagcgtacgac.....ccggttgagaccgacacat 208

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues  
Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :	N_Geneseq_101002.*			
	1:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*		
	2:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*		
	3:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*		
	4:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*		
	5:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*		
	6:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*		
	7:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*		
	8:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*		
	9:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*		
	10:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*		
	11:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*		
	12:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*		
	13:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*		
	14:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*		
	15:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*		
	16:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*		
	17:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*		
	18:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*		
	19:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*		
	20:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*		
	21:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*		
	22:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*		
	23:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*		
	24:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*		

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	208	100.0	208	22	AAS05201 Mycobacterium gord
2	195.2	93.8	208	22	AAS05203 Mycobacterium gord
3	190.4	91.5	208	22	AAS05215 Mycobacterium avi
4	177.6	85.4	208	22	AAS05202 Mycobacterium gord
5	177.6	85.4	208	22	AAS05211 Mycobacterium mari
6	175.2	84.2	207	22	AAS05204 Mycobacterium gord
7	172.8	83.1	208	22	AAS05210 Mycobacterium ulce
8	169.6	81.5	208	22	AAS05219 Mycobacterium fort
9	168	80.8	205	22	AAS05220 Mycobacterium intr

10	168	80.8	208	22	AAS05208 Mycobacterium kans
11	166.4	80.0	208	22	AAS05205 Mycobacterium tube
12	166.4	80.0	208	22	AAS05206 Mycobacterium terr
13	166.4	80.0	208	22	AAS05216 Mycobacterium bovi
14	166.4	80.0	208	22	AAS05217 Mycobacterium ceta
15	166.4	80.0	970	17	AAT09676 Mycobacterium tube
16	166.4	80.0	3519	22	AAH51976 Mycobacterium tube
17	166.4	80.0	3534	22	AAH02079 Mycobacterium tube
18	166.4	80.0	3853	21	AAA74651 Mycobacterium tube
19	166.4	80.0	3853	21	AAA89994 M. tuberculosis rp
20	164.8	79.2	208	22	AAS05222 Mycobacterium afri
21	164.8	79.2	208	22	AAS05223 Mycobacterium haem
22	164	78.8	207	22	AAS05212 Mycobacterium szul
23	155.2	74.6	208	22	AAS05224 Mycobacterium xeno
24	142.4	68.5	214	22	AAS05221 Mycobacterium abs
25	142.4	68.5	3447	14	AAH51532 M.leprae rpoB gene
26	140.8	67.7	214	22	AAS05214 Mycobacterium malm
27	139.8	67.2	223	22	AAS05209 Mycobacterium scro
28	139.2	66.9	214	22	AAS05207 Mycobacterium chel
29	135	64.9	223	22	AAS05213 Mycobacterium gast
30	131	63.0	211	22	AAS05218 Mycobacterium flav
31	83.2	40.0	3495	22	AAH65512 C glutamicum codin
32	83.2	40.0	34980	22	AAH68525 C glutamicum codin
33	73.6	35.4	27426	23	AAS59541 Propionibacterium
34	49.8	23.9	6085	18	AAT70153 S. longisporoflavus
c 35	45.4	21.8	4403765	22	AA199683 Mycobacterium tube
36	45.2	21.7	2104	21	AAAC75978 Human ORFX ORF1533
37	44	21.2	21185	21	AAAG3350 Streptomyces globi
38	44	21.2	63164	21	AAAG3348 Streptomyces globi
39	43	20.7	2122	18	AAT70152 S. longisporoflavus
c 40	42.2	20.3	77536	21	AAA14651 Nucleotide sequenc
41	42	20.2	1844	12	AAQ12000 Maize nitrite redu
42	42	20.2	2244	17	AAT10429 Mature isoamylase
43	42	20.2	2244	19	AAV23640 Flavobacterium iso
44	42	20.2	2334	17	AAT10428 Isoamylase gene.
45	42	20.2	2334	19	AAV23639 Full length Flavob

ALIGNMENTS

RESULT 1

AAS05201 ID AAS05201 standard; DNA; 208 BP.

AC AAS05201;

DT 07-SEP-2001 (first entry)

DE Mycobacterium gordonae type I rpoB gene fragment.

```
XX Claim 1; Page 40; 50pp; English.
XX
XX The present sequence for Mycobacterium gordonae type I rpoB gene
CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
CC various Mycobacterial species. These rpoB gene fragments can be used
CC in the diagnosis and identification of Mycobacterium species using a
CC novel PCR-restriction fragment length polymorphism analysis (PRA)
CC method. The method comprises obtaining a restriction fragment length
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
CC amplifying and digesting the DNA fragment from the microorganism to
CC be identified and comparing the RFLP patterns from the known rpoB gene
CC fragments with the unidentified fragment. The rpoB gene fragments
CC are useful to identify a wide range of Mycobacterium species, e.g. for
CC diagnosis or to obtain epidemiological and pathogenesis information for
CC selection of appropriate therapies, including M. tuberculosis, M. leprae
CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene
CC fragments is rapid, precise, simple and cost effective (only 1 PCR
CC required), and can differentiate between many species in a single
CC experiment, including those difficult to distinguish by usual biochemical
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
CC detecting specific Mycobacterial species.
XX
XX Sequence 208 BP; 46 A; 71 C; 65 G; 26 T; 0 other;
XX
XX Query Match 100.0%; Score 208; DB 22; Length 208;
XX Best Local Similarity 100.0%; Pred. No. 8.9e-36;
XX Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 TCAAGAGGAAGCGCTACGACCTGGCCGGGTAGGCGCTACAAGGTCAACAAGAAGCTCG 60
DB 1 TCAGGAGAAGCGCTACGACCTGGCCGGGTAGGCGCTACAAGGTCAACAAGAAGCTCG 60
QY 61 GCCTGACCTCGCGGATCGGATCACCAGCTGCACGCTGACCGAGGAAGACGTCGTCGCCA 120
DB 61 GCCTGACCTCGCGGATCGGATCACCAGCTGCACGCTGACCGAGGAAGACGTCGTCGCCA 120
QY 121 CCATCGAGTACCTGGTTCGCGCTGCAGGGCCAGCACACGATGACCGTCCCGGGCGGCA 180
DB 121 CCATCGAGTACCTGGTTCGCGCTGCAGGGCCAGCACACGATGACCGTCCCGGGCGGCA 180
QY 181 CCGAGGTGCGCGTTGAGACCGACGACAT 208
DB 181 CCGAGGTGCGCGTTGAGACCGACGACAT 208
XX
XX AAS05203;
XX
XX 07-SEP-2001 (first entry)
XX
XX Mycobacterium gordonae type III rpoB gene fragment.
XX
XX Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;
XX PCR-restriction fragment length polymorphism analysis; ds.
XX
XX Mycobacterium gordonae type III.
XX
XX WO200131061-A1.
XX
XX 03-MAY-2001.
XX
XX 27-OCT-2000; 2000WO-KR01223.
XX
XX 27-OCT-1999; 99KR-0046795.
XX
XX (ERUM-) ERUME BIOTECH CO LTD.
XX
XX Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;
XX
```

```
XX WPI; 2001-300520/31.
XX
XX New DNA fragments from the rpoB gene of mycobacteria, useful for
PT diagnosis and identification of many mycobacterial species by
PT restriction fragment length polymorphism -
XX
XX Claim 1; Page 41; 50pp; English.
XX
XX The present sequence for Mycobacterium gordonae type III rpoB gene
CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
CC various Mycobacterial species. These rpoB gene fragments can be used
CC in the diagnosis and identification of Mycobacterium species using a
CC novel PCR-restriction fragment length polymorphism analysis (PRA)
CC method. The method comprises obtaining a restriction fragment length
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
CC amplifying and digesting the DNA fragment from the microorganism to
CC be identified and comparing the RFLP patterns from the known rpoB gene
CC fragments with the unidentified fragment. The rpoB gene fragments
CC are useful to identify a wide range of Mycobacterium species, e.g. for
CC diagnosis or to obtain epidemiological and pathogenesis information for
CC selection of appropriate therapies, including M. tuberculosis, M. leprae
CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene
CC fragments is rapid, precise, simple and cost effective (only 1 PCR
CC required), and can differentiate between many species in a single
CC experiment, including those difficult to distinguish by usual biochemical
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
CC detecting specific Mycobacterial species.
XX
XX Sequence 208 BP; 46 A; 69 C; 63 G; 30 T; 0 other;
XX
XX Query Match 93.8%; Score 195.2; DB 22; Length 208;
XX Best Local Similarity 96.2%; Pred. No. 4.7e-33;
XX Matches 200; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
XX
QY 1 TCAAGGAGAAGCGCTACGACCTGGCCGGGTAGGCGCTACAAGGTCAACAAGAAGCTCG 60
DB 1 TCAAGGAGAAGCGCTACGACCTGGCCGGGTAGGCGCTACAAGGTCAACAAGAAGCTCG 60
QY 61 GCCTGACCTCGCGGATCGGATCACCAGCTGCACGCTGACCGAGGAAGACGTCGTCGCCA 120
DB 61 GCCTGACCTCGCGGATCGGATCACCAGCTGCACGCTGACCGAGGAAGACGTCGTCGCCA 120
QY 121 CCATCGAGTACCTGGTTCGCGCTGCAGGGCCAGCACACGATGACCGTCCCGGGCGGCA 180
DB 121 CCATCGAGTACCTGGTTCGCGCTGCAGGGCCAGCACACGATGACCGTCCCGGGCGGCA 180
QY 181 CCGAGGTGCGCGTTGAGACCGACGACAT 208
DB 181 CCGAGGTGCGCGTTGAGACCGACGACAT 208
XX
XX RESULT 3
XX AAS05215
XX ID AAS05215 standard; DNA; 208 BP.
XX
XX AAS05215;
XX
XX 07-SEP-2001 (first entry)
XX
XX Mycobacterium avium rpoB gene fragment.
XX
XX Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;
XX PCR-restriction fragment length polymorphism analysis; ds.
XX
XX Mycobacterium avium.
XX
XX WO200131061-A1.
XX
XX 03-MAY-2001.
XX
XX 27-OCT-2000; 2000WO-KR01223.
XX
XX
```

```
XX 27-OCT-1999; 99KR-0046795.
XX (ERUM-) ERUME BIOTECH CO LTD.
XX Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;
XX WPI; 2001-300520/31.
XX
XX New DNA fragments from the rpoB gene of mycobacteria, useful for
XX diagnosis and identification of many mycobacterial species by
XX restriction fragment length polymorphism -
XX
XX Claim 1; Page 44; 50pp; English.
XX
XX The present sequence for Mycobacterium avium rpoB gene
XX fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
XX various Mycobacterial species. These rpoB gene fragments can be used
XX in the diagnosis and identification of Mycobacterium species using a
XX novel PCR-restriction fragment length polymorphism analysis (PRA)
XX method. The method comprises obtaining a restriction fragment length
XX polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
XX amplifying and digesting the DNA fragment from the microorganism to
XX be identified and comparing the RFLP patterns from the known rpoB gene
XX fragments with the unidentified fragment. The rpoB gene fragments
XX are useful to identify a wide range of Mycobacterium species, e.g. for
XX diagnosis or to obtain epidemiological and pathogenesis information for
XX selection of appropriate therapies, including M. tuberculosis, M. leprae
XX and non-tuberculous mycobacteria (NTM) encountered in subjects infected
XX with human immunodeficiency virus (HIV). Analysis of the rpoB gene
XX fragments is rapid, precise, simple and cost effective (only 1 PCR
XX required), and can differentiate between many species in a single
XX experiment, including those difficult to distinguish by usual biochemical
XX tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
XX detecting specific Mycobacterial species.
XX
XX Sequence 208 BP; 44 A; 69 C; 69 G; 26 T; 0 other;
XX
XX Query Match 91.5%; Score 190.4; DB 22; Length 208;
XX Best Local Similarity 94.7%; Pred. No. 5e-32;
XX Matches 197; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
XX
XX 1 TCAAGGAGAGCGCTACGACCTGCGCCGGTAGCCGCTACAAGGTCAACAAGAGCTCG 60
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 1 TCAAGGAGAGCGCTACGACCTGCGCCGGTAGCCGCTACAAGGTCAACAAGAGCTCG 60
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 61 GCCTGCAGCTCGCGATCCGATCACAGCTCCACGCTGACCGAGGAGAGCTGTCGCCA 120
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 61 GCCTGCAGCTCGCGATCCGATCACAGCTCCACGCTGACCGAGGAGAGCTGTCGCCA 120
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 121 CCATCGAGTACCTGCTCGCTGCAGGAGGCCAGCACAGCATGACCGTCCCGGGCGGCA 180
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 121 CCATCGAGTACCTGCTCGCTGCAGGAGGCCAGCACAGCATGACCGTCCCGGGCGGCA 180
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 181 CCGAGGTGCGGGTTGAGACCGACGACAT 208
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 181 TCGAGGTGCGGGTTGAGACCGACGACAT 208
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX
XX RESULT 4
XX AAS05202
XX ID AAS05202 standard; DNA; 208 BP.
XX XX
XX AC AAS05202;
XX XX
XX DT 07-SEP-2001 (first entry)
XX XX
XX DE Mycobacterium gordonae type II rpoB gene fragment.
XX XX
XX KW Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;
XX PCR-restriction fragment length polymorphism analysis; ds.
XX XX
XX OS Mycobacterium gordonae type II.
```

```
XX WO200131061-A1.
XX 03-MAY-2001.
XX 27-OCT-2000; 2000WO-KR01223.
XX 27-OCT-1999; 99KR-0046795.
XX (ERUM-) ERUME BIOTECH CO LTD.
XX Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;
XX WPI; 2001-300520/31.
XX
XX New DNA fragments from the rpoB gene of mycobacteria, useful for
XX diagnosis and identification of many mycobacterial species by
XX restriction fragment length polymorphism -
XX
XX Claim 1; Page 40; 50pp; English.
XX
XX The present sequence for Mycobacterium gordonae type II rpoB gene
XX fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
XX various Mycobacterial species. These rpoB gene fragments can be used
XX in the diagnosis and identification of Mycobacterium species using a
XX novel PCR-restriction fragment length polymorphism analysis (PRA)
XX method. The method comprises obtaining a restriction fragment length
XX polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
XX amplifying and digesting the DNA fragment from the microorganism to
XX be identified and comparing the RFLP patterns from the known rpoB gene
XX fragments with the unidentified fragment. The rpoB gene fragments
XX are useful to identify a wide range of Mycobacterium species, e.g. for
XX diagnosis or to obtain epidemiological and pathogenesis information for
XX selection of appropriate therapies, including M. tuberculosis, M. leprae
XX and non-tuberculous mycobacteria (NTM) encountered in subjects infected
XX with human immunodeficiency virus (HIV). Analysis of the rpoB gene
XX fragments is rapid, precise, simple and cost effective (only 1 PCR
XX required), and can differentiate between many species in a single
XX experiment, including those difficult to distinguish by usual biochemical
XX tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
XX detecting specific Mycobacterial species.
XX
XX Sequence 208 BP; 45 A; 62 C; 73 G; 28 T; 0 other;
XX
XX Query Match 85.4%; Score 177.6; DB 22; Length 208;
XX Best Local Similarity 90.9%; Pred. No. 2.7e-29;
XX Matches 189; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
XX
XX 1 TCAAGGAGAGCGCTACGACCTGCGCCGGTAGCCGCTACAAGGTCAACAAGAGCTCG 60
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 1 TCAAGGAGAGCGCTACGACCTGCGCCGGTAGCCGCTACAAGGTCAACAAGAGCTCG 60
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 61 GCCTGCAGCTCGCGATCCGATCACAGCTCCACGCTGACCGAGGAGAGCTGTCGCCA 120
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 61 GTCTGAACGTGCGGAGCGGATCCAGCTCCAGCTGACCGAGGAGAGCTGTCGTAGCCA 120
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 121 CCATCGAGTACCTGCTCGCTGCAGGAGGCCAGCACAGCATGACCGTCCCGGGCGGCA 180
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 121 CCATCGAGTACCTGCTCGCTGCAGGAGGCCAGCACAGCATGACCGTCCCGGGCGGCG 180
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 181 CCGAGGTGCGGGTTGAGACCGGACGACAT 208
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 181 CCGAGGTGCGGGTTGAGACCGGACGACAT 208
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX
XX RESULT 5
XX AAS05211
XX ID AAS05211 standard; DNA; 208 BP.
XX XX
XX AC AAS05211;
XX XX
XX DT 07-SEP-2001 (first entry)
XX XX
```

DE Mycobacterium marinum rpoB gene fragment.  
XX  
KW Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;  
KW PCR-restriction fragment length polymorphism analysis; ds.  
XX  
XX Mycobacterium marinum.  
XX  
XX WO200131061-A1.  
XX  
XX 03-MAY-2001.  
XX  
XX 27-OCT-2000; 2000WO-KR01223.  
XX  
XX 27-OCT-1999; 99KR-0046795.  
XX  
XX (ERUM-) ERUME BIOTECH CO LTD.  
XX  
XX Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;  
XX WPI: 2001-300520/31.  
XX  
XX New DNA fragments from the rpoB gene of mycobacteria, useful for  
PT diagnosis and identification of many mycobacterial species by  
PT restriction fragment length polymorphism -  
XX  
XX Claim 1; Page 43; 50pp; English.  
XX  
XX The present sequence for Mycobacterium marinum rpoB gene  
CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from  
CC various Mycobacterial species. These rpoB gene fragments can be used  
CC in the diagnosis and identification of Mycobacterium species using a  
CC novel PCR-restriction fragment length polymorphism analysis (PRA)  
CC method. The method comprises obtaining a restriction fragment length  
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,  
CC amplifying and digesting the DNA fragment from the microorganism to  
CC be identified and comparing the RFLP patterns from the known rpoB gene  
CC fragments with the unidentified fragment. The rpoB gene fragments  
CC are useful to identify a wide range of Mycobacterium species, e.g. for  
CC diagnosis or to obtain epidemiological and pathogenesis information for  
CC selection of appropriate therapies, including M. tuberculosis, M. leprae  
CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected  
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene  
CC fragments is rapid, precise, simple and cost effective (only 1 PCR  
CC required), and can differentiate between many species in a single  
CC experiment, including those difficult to distinguish by usual biochemical  
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for  
CC detecting specific Mycobacterial species.  
XX  
SQ Sequence 208 BP; 45 A; 69 C; 68 G; 26 T; 0 other;  
Query Match 85.4%; Score 177.6; DB 22; Length 208;  
Best Local Similarity 90.9%; Pred. No. 2.7e-29;  
Matches 189; Conservative 0; Mismatches 19; Indels 0; Gaps 0;  
QY 1 TCAAGGAGAGCGCTACGACCTGGCCGGGTAGGCGCTACAAGGTCAACAAGAGCTCG 60  
Db 1 TCAAGGAGAGCGCTACGACCTGGCCGGGTAGGCGCTACAAGGTCAACAAGAGCTCG 60  
QY 61 GCCTGACGCTGGCGATCCGATCACCAGCTCCACGCTGACCGAGGAGACGCTGTCGCCA 120  
Db 61 GCCTGAACCGCGCCAGCCATCAGCAGCTCGACGCTGACCGAGGAGACGCTGTCGCCA 120  
QY 121 CCATCGAGTACCTGGTCCGCTGCACGAGGCGCCACACACGATGACCGTCCGGCGGCA 180  
Db 121 CCATGAATACCTGGTCCGCTTGCACGAGGCGCCAGACCGGATGACCGTCCGGCGGTG 180  
QY 181 CCGAGTGCCGGTTGAGACCGACGACAT 208  
Db 181 TCGAGTGCCGGTTCGAGACCGACGACAT 208  
RESULT 6  
AAS05204

ID AAS05204 standard; DNA; 207 BP.  
XX  
AC AAS05204;  
XX  
DT 07-SEP-2001 (first entry)  
XX  
XX Mycobacterium gordonae type IV rpoB gene fragment.  
XX  
XX WO200131061-A1.  
XX  
XX 03-MAY-2001.  
XX  
XX 27-OCT-2000; 2000WO-KR01223.  
XX  
XX 27-OCT-1999; 99KR-0046795.  
XX  
XX (ERUM-) ERUME BIOTECH CO LTD.  
XX  
XX Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;  
XX WPI: 2001-300520/31.  
XX  
XX New DNA fragments from the rpoB gene of mycobacteria, useful for  
PT diagnosis and identification of many mycobacterial species by  
PT restriction fragment length polymorphism -  
XX  
XX Claim 1; Page 41; 50pp; English.  
XX  
XX The present sequence for Mycobacterium gordonae type IV rpoB gene  
CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from  
CC various Mycobacterial species. These rpoB gene fragments can be used  
CC in the diagnosis and identification of Mycobacterium species using a  
CC novel PCR-restriction fragment length polymorphism analysis (PRA)  
CC method. The method comprises obtaining a restriction fragment length  
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,  
CC amplifying and digesting the DNA fragment from the microorganism to  
CC be identified and comparing the RFLP patterns from the known rpoB gene  
CC fragments with the unidentified fragment. The rpoB gene fragments  
CC are useful to identify a wide range of Mycobacterium species, e.g. for  
CC diagnosis or to obtain epidemiological and pathogenesis information for  
CC selection of appropriate therapies, including M. tuberculosis, M. leprae  
CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected  
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene  
CC fragments is rapid, precise, simple and cost effective (only 1 PCR  
CC required), and can differentiate between many species in a single  
CC experiment, including those difficult to distinguish by usual biochemical  
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for  
CC detecting specific Mycobacterial species.  
XX  
SQ Sequence 207 BP; 45 A; 66 C; 66 G; 30 T; 0 other;  
Query Match 84.2%; Score 175.2; DB 22; Length 207;  
Best Local Similarity 93.3%; Pred. No. 8.7e-29;  
Matches 194; Conservative 0; Mismatches 13; Indels 1; Gaps 1;  
QY 1 TCAAGGAGAGCGCTACGACCTGGCCGGGTAGGCGCTACAAGGTCAACAAGAGCTCG 60  
Db 1 TCAAGGAGAGCGCTACGACCTGGCCGGGTAGGCGCTACAAGGTCAACAAGAGCTCG 60  
QY 61 GCCTGACGCTGGCGATCCGATCACCAGCTCCACGCTGACCGAGGAGACGCTGTCGCCA 120  
Db 61 GCCTGTCATGTCGGCGATCCGATCACCAGCTGACCGTACCGAAGAGGAGCTGTCGCCA 120  
QY 121 CCATCGAGTACCTGGTCCGCTGCACGAGGCGCCACACACGATGACCGTCCGGCGGCA 180  
Db 121 CCATCGAGTACCTGGTCCGCTTCCACGAGGCGCCAGACCGATGACCGTCCGGCGGGA 179  
QY 181 CCGAGGTGCCGGTTGAGACCGACGACAT 208

Db	180	CCGAGGTTCCGGTGGAGACCGACGACAT	207	
RESULT 7				
AA505210				
ID	AA505210	standard; DNA; 208 BP.		
XX	AA505210;			
XX	07-SEP-2001	(first entry)		
XX	Mycobacterium ulcerans	rpob gene fragment.		
XX	Non-tuberculous mycobacteria;	rpob gene fragment; NTM; HIV; PRA; RFLP;		
XX	PCR-restriction fragment length polymorphism analysis;	ds.		
XX	Mycobacterium ulcerans.			
OS	Mycobacterium ulcerans.			
XX	WO200131061-A1.			
PN	03-MAY-2001.			
XX	27-OCT-2000;	2000WO-KR01223.		
PF	27-OCT-1999;	99KR-0046795.		
XX	(ERUN-)	ERUME BIOTECH CO LTD.		
PA	Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;			
PI	WPI; 2001-300520/31.			
XX	New DNA fragments from the rpob gene of mycobacteria, useful for			
XX	diagnosis and identification of many mycobacterial species by			
PT	restriction fragment length polymorphism -			
PT	Claim 1; Page 43; 50pp; English.			
XX	The present sequence for Mycobacterium ulcerans rpob gene			
XX	fragment is 1 of 24 rpob gene fragments (AA505201-AA505224) from			
CC	various Mycobacterial species. These rpob gene fragments can be used			
CC	in the diagnosis and identification of Mycobacterium species using a			
CC	novel PCR-restriction fragment length polymorphism analysis (PRA)			
CC	method. The method comprises obtaining a restriction fragment length			
CC	polymorphism (RFLP) pattern of the 24 rpob gene fragments; isolating,			
CC	amplifying and digesting the DNA fragment from the microorganism to			
CC	be identified and comparing the RFLP patterns from the known rpob gene			
CC	fragments with the unidentified fragment. The rpob gene fragments			
CC	are useful to identify a wide range of Mycobacterium species, e.g. for			
CC	diagnosis or to obtain epidemiological and pathogenesis information for			
CC	selection of appropriate therapies, including M. tuberculosis, M. leprae			
CC	and non-tuberculous mycobacteria (NTM) encountered in subjects infected			
CC	with human immunodeficiency virus (HIV). Analysis of the rpob gene			
CC	fragments is rapid, precise, simple and cost effective (only 1 PCR			
CC	required), and can differentiate between many species in a single			
CC	experiment, including those difficult to distinguish by usual biochemical			
CC	tests. Also described are oligonucleotide probes (AA505227-AA505242) for			
CC	detecting specific Mycobacterial species.			
XX	Sequence 208 BP; 45 A; 68 C; 67 G; 28 T; 0 other;			
XX	Query Match	83.1%; Score 172.8; DB 22; Length 208;		
XX	Best Local Similarity	89.4%; Pred. No. 2.8e-28;		
XX	Matches 196; Conservative	0; Mismatches 22; Indels 0; Gaps 0;		
OY	1	TCAAGGAGAAGCGCTACGACCTGGCGGGTAGGCGCGCTACAAGGTCAACAAGAAGCTCG	60	
DB	1	TCAAGGAGAAGCGCTACGACCTGGCTGCGTGGTGGTACAGGTCAACAAGAAGCTCG	60	
OY	61	GCCTGCACGTGGCGGATTCGATCACCACTCTACCGTGTACCGAGGAAGAGCTGCTGCCA	120	
DB	61	GCCTGAACGCCGCCAGCCATCACCACTGACGCTGACGAGGAAGAGCTGCTGCCA	120	

Qy	121	CCATCGAGTACCTGGTCCGCTGCACGAGGCCAGCACACGATGACCGTCCCGGGCGGCA	180
Db	121	CTCGAATACTCTGGTCCGCTTGACGAGGCCAGACCGGATGACCGTCCGGGGCGTG	180
Qy	181	CCGAGGTGCCGGTGTGACCGAGCAGAT	208
Db	181	TCGAGGTGCCGGTGTGACCGAGCAGCAGAT	208
RESULT 8			
AAS05219			
ID	AAS05219	standard; DNA; 208 BP.	
XX			
AC	AAS05219;		
XX			
DT	07-SEP-2001	(first entry)	
XX			
DE	Mycobacterium fortuitum rpoB gene fragment.		
XX			
KW	Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP		
KW	PCR-restriction fragment length polymorphism analysis; ds.		
XX			
OS	Mycobacterium fortuitum.		
XX			
PN	W0200131061-AL.		
XX			
PD	03-MAY-2001.		
XX			
XX	27-OCT-2000; 2000MO-KR01223.		
PF			
XX			
PR	27-OCT-1999; 99KR-0046795.		
XX			
PA	(ERUM-) ERUME BIOTECH CO LTD.		
XX			
PI	Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;		
XX			
DR	WPI; 2001-300520/31.		
XX			
PT	New DNA fragments from the rpoB gene of mycobacteria, useful for		
PT	diagnosis and identification of many mycobacterial species by		
PT	restriction fragment length polymorphism -		
XX			
PS	Claim 1; Page 46; 50pp; English.		
XX			
CC	The present sequence for Mycobacterium fortuitum rpoB gene		
CC	fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from		
CC	various Mycobacterial species. These rpoB gene fragments can be used		
CC	in the diagnosis and identification of Mycobacterium species using a		
CC	novel PCR-restriction fragment length polymorphism analysis (PRA)		
CC	method. The method comprises obtaining a restriction fragment length		
CC	polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,		
CC	amplifying and digesting the DNA fragment from the microorganism to		
CC	be identified and comparing the RFLP patterns from the known rpoB		
CC	fragments with the unidentified fragment. The rpoB gene fragments		
CC	are useful to identify a wide range of Mycobacterium species, e.g. for		
CC	diagnosis or to obtain epidemiological and pathogenesis information for		
CC	selection of appropriate therapies, including M. tuberculosis, M. lepre		
CC	and non-tuberculous mycobacteria (NTM) encountered in subjects infect		
CC	with human immunodeficiency virus (HIV). Analysis of the rpoB gene		
CC	fragments is rapid, precise, simple and cost effective (only 1 PCR		
CC	required), and can differentiate between many species in a single		
CC	experiment, including those difficult to distinguish by usual biochem		
CC	tests. Also described are oligonucleotide probes (AAS05227-AAS05242)		
CC	detecting specific Mycobacterial species.		
XX			
SQ	Sequence 208 BP; 43 A; 69 C; 69 G; 27 T; 0 other;		
Query Match	81.5%;	Score 169.6;	DB 22; Length 208;
Best Local Similarity	88.5%;	Pred. No. 1.4e-27;	
Matches 184;	Conservative 0;	Mismatches 24;	Indels 0; Gaps
Qy	1	TCAAGGAGACGCGTACGACCTGGCCGGGTAGCGCGCTACAAGGTCAACAGAAGCTCG	60

```

Db      1 TCAAGAGAGCGCTACGACCTGGCCGCTGAGCGCTACAAAGTCAACAAGAGCTGG 60
Qy      61 GCCTGCACGTCGGGATCCAGATCAGAGCTGCACGCTGACCGAGGAAGACGTCGTCGCCA 120
Db      61 GCCTGAACGCCGCCAGCGGATCAGCTCGTCTGACCTGACCGAGGAAGACGTCGTCGCCA 120
Qy      121 CCATCCAGTACCTGGTCCGCTGCAGGAGGCCAGCACACGATGACCGTCCCGGCGGCCA 180
Db      121 CCATCCAGTACCTGGTCCGCTGCAGGAGGCCAGCACACGATGACCGTCCCGGCGGCCA 180
Qy      181 CCGAGGTCCCGTGTGAGCCGACGACAT 208
Db      181 TCGAGTCCCGTGTGAGGTCGAGACAT 208

RESULT 9
AAS05220
ID AAS05220 standard; DNA; 205 BP.
XX
AC AAS05220;
XX
DT 07-SEP-2001 (first entry)
XX
DE Mycobacterium intracellulare rpoB gene fragment.
XX
KW Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;
KW PCR-restriction fragment length polymorphism analysis; ds.
XX
OS Mycobacterium intracellulare.
XX
PN WO200131061-A1.
XX
PD 03-MAY-2001.
XX
PF 27-OCT-2000; 2000WO-KR01223.
XX
PR 27-OCT-1999; 99KR-0046795.
XX
PA (ERUM-) ERUME BIOTECH CO LTD.
XX
PI Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;
XX
PS WPI; 2001-300520/31.
XX
CC The present sequence for Mycobacterium intracellulare rpoB gene
CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
CC various Mycobacterial species. These rpoB gene fragments can be used
CC in the diagnosis and identification of Mycobacterium species using a
CC novel PCR-restriction fragment length polymorphism analysis (PRA)
CC method. The method comprises obtaining a restriction fragment length
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
CC amplifying and digesting the DNA fragment from the microorganism to
CC be identified and comparing the RFLP patterns from the known rpoB gene
CC fragments with the unidentified fragment. The rpoB gene fragments
CC are useful to identify a wide range of Mycobacterium species, e.g. for
CC diagnosis or to obtain epidemiological and pathogenesis information for
CC selection of appropriate therapies, including M. tuberculosis, M. leprae
CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene
CC fragments is rapid, precise, simple and cost effective (only 1 PCR
CC required), and can differentiate between many species in a single
CC experiment, including those difficult to distinguish by usual biochemical
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
CC detecting specific Mycobacterial species.
XX
SQ Sequence 205 BP; 44 A; 69 C; 67 G; 25 T; 0 other;

```

```

Query Match      80.8%; Score 168; DB 22; Length 205;
Best Local Similarity 91.3%; Pred. No. 3e-27;
Matches 190; Conservative 0; Mismatches 15; Indels 3; Gaps 1;

Qy      1 TCAAGAGAGAGCGCTACGACCTGGCCCGGGTAGCCGCTACAAAGTCAACAAGAGCTCG 60
Db      1 TCAAGAGAGAGCGCTACGACCTGGCGCTGTCGGCCGCTACAAAGTCAACAAGAGCTCG 60
Qy      61 GCCTGCACGTCGGGATCCGATCACCAGCTCCAGCTGACCGAGGAGACGTCGTCGCCA 120
Db      61 GCCTGCACGTCGGGATCCGATCACCAGCTCCAGCTGACCGAGGAGACGTCGTCGCCA 120
Qy      121 CCATCGAGTACCTGGTCCGCTGCACGAGGCCAGCACACGATGACCGTCCCGGCGGCCA 180
Db      121 CCATCGAGTACCTGGTCCGCTGCACGAGGCCAGCCACGATGACCGTCCCGGCGGCCA 180
Qy      181 CCGAGGTCCCGTGTGACACCGAGACAT 208
Db      178 TCGAGGTCCCGTGTGAGACCGAGACAT 205

RESULT 10
AAS05208
ID AAS05208 standard; DNA; 208 BP.
XX
AC AAS05208;
XX
DT 07-SEP-2001 (first entry)
XX
DE Mycobacterium kansasii rpoB gene fragment.
XX
KW Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;
KW PCR-restriction fragment length polymorphism analysis; ds.
XX
OS Mycobacterium kansasii.
XX
PN WO200131061-A1.
XX
PD 03-MAY-2001.
XX
PF 27-OCT-2000; 2000WO-KR01223.
XX
PR 27-OCT-1999; 99KR-0046795.
XX
PA (ERUM-) ERUME BIOTECH CO LTD.
XX
PI Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;
XX
PS WPI; 2001-300520/31.
XX
CC New DNA fragments from the rpoB gene of mycobacteria, useful for
CC diagnosis and identification of many mycobacterial species by
CC restriction fragment length polymorphism -
CC
CC Claim 1; Page 42; 50pp; English.
CC
CC The present sequence for Mycobacterium kansasii rpoB gene
CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
CC various Mycobacterial species. These rpoB gene fragments can be used
CC in the diagnosis and identification of Mycobacterium species using a
CC novel PCR-restriction fragment length polymorphism analysis (PRA)
CC method. The method comprises obtaining a restriction fragment length
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
CC amplifying and digesting the DNA fragment from the microorganism to
CC be identified and comparing the RFLP patterns from the known rpoB gene
CC fragments with the unidentified fragment. The rpoB gene fragments
CC are useful to identify a wide range of Mycobacterium species, e.g. for
CC diagnosis or to obtain epidemiological and pathogenesis information for
CC selection of appropriate therapies, including M. tuberculosis, M. leprae
CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene
CC fragments is rapid, precise, simple and cost effective (only 1 PCR

```

CC required), and can differentiate between many species in a single  
CC experiment, including those difficult to distinguish by usual biochemical  
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for  
CC detecting specific Mycobacterial species.

XX  
SQ Sequence 208 BP; 51 A; 65 C; 65 G; 27 T; 0 other;

Query Match 80.8%; Score 168; DB 22; Length 208;

Best Local Similarity 88.0%; Pred. No. 3e-27; Length 208;  
Matches 183; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY 1 TCAAGGAGAACGCTACGACCTGGCCCGGTAGCCGCTACAGGTCAACAAGAGCTCG 60

|||||  
Db 1 TCAAGGAGAACGCTACGACCTGGCCCGGTAGCCGCTACAGGTCAACAAGAGCTCG 60

QY 61 GCCTGCACGTGCGCGATCCGATCCACGCTCCACGCTGACCGAGGAAGACGCTCGTGC 120

|||||  
Db 61 GCCTGAACCAATATCCGATCCACGACGACGCTGACCGAAGAAGACGCTCGTGC 120

QY 121 CCATCGATACCTGCTCCGCTGCACGAGGCGCAGACGATGACCGTCCCGGGCGGCA 180

|||||  
Db 121 CCATCGATATCTGCTCCGCTGCACGAGGCGCAGGCGCAGGCGTCCCGGGCGG 180

QY 181 CCGAGGTGCGGCTTGAGACCGAGCAT 208

|||||  
Db 181 TCGAGGTGCGGCTGGAACCGAGCAT 208

|||||

RESULT 11

AAS05205

ID AAS05205 standard; DNA: 208 BP.

XX AAS05205;

AC AAS05205;

DT 07-SEP-2001 (first entry)

XX Mycobacterium tuberculosis rpoB gene fragment.

DE Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;

KW PCR-restriction fragment length polymorphism analysis; ds.

XX Mycobacterium tuberculosis.

OS WO200131061-A1.

PN 03-MAY-2001.

XX 27-OCT-2000; 2000WO-KR01223.

XX 27-OCT-1999; 99KR-0046795.

PR (ERUM-) ERUME BIOTECH CO LTD.

PA Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;

XX WPI; 2001-300520/31.

XX New DNA fragments from the rpoB gene of mycobacteria, useful for

PT diagnosis and identification of many mycobacterial species by

PT restriction fragment length polymorphism -

PS Disclosure; Page 41; 50pp; English.

XX The present sequence for Mycobacterium tuberculosis rpoB gene

CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from

CC various Mycobacterial species. These rpoB gene fragments can be used

CC in the diagnosis and identification of Mycobacterium species using a

CC novel PCR-restriction fragment length polymorphism analysis (PRA)

CC method. The method comprises obtaining a restriction fragment length

CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,

CC amplifying and digesting the DNA fragment from the microorganism to

CC be identified and comparing the RFLP patterns from the known rpoB gene

CC fragments with the unidentified fragment. The rpoB gene fragments

CC are useful to identify a wide range of Mycobacterium species, e.g. for  
CC diagnosis or to obtain epidemiological and pathogenesis information for  
CC selection of appropriate therapies, including M. tuberculosis, M. leprae  
CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected  
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene  
CC fragments is rapid, precise, simple and cost effective (only 1 PCR  
CC required), and can differentiate between many species in a single  
CC experiment, including those difficult to distinguish by usual biochemical  
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for  
CC detecting specific Mycobacterial species.

SQ Sequence 208 BP; 46 A; 62 C; 67 G; 33 T; 0 other;

Query Match 80.0%; Score 166.4; DB 22; Length 208;

Best Local Similarity 87.5%; Pred. No. 6.5e-27;

Matches 182; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 1 TCAAGGAGAACGCTACGACCTGGCCCGGTAGCCGCTACAGGTCAACAAGAGCTCG 60

|||||  
Db 1 TCAAGGAGAACGCTACGACCTGGCCCGGTAGCCGCTACAGGTCAACAAGAGCTCG 60

QY 61 GCCTGCACGTGCGCGATCCGATCCACGCTCCACGCTGACCGAGGAAGACGCTCGTGC 120

|||||  
Db 61 GCCTGCATGTCGCGAGCCCATCAGCTCGTGCAGCTGACCGAAGAAGACGCTCGTGC 120

QY 121 CCATCGATACCTGCTCGGCTGCACGAGGCGCAGACGATGACCGTCCCGGGCGGCA 180

|||||  
Db 121 CCATCGATATCTGCTCGGCTGCACGAGGCGCAGACGATGACCGTCCCGGGCGG 180

QY 181 CCGAGGTGCGGCTTGAGACCGAGCAT 208

|||||  
Db 181 TCGAGGTGCGGCTGGAACCGAGCAT 208

|||||

RESULT 12

AAS05206

ID AAS05206 standard; DNA: 208 BP.

XX AAS05206;

AC AAS05206;

DT 07-SEP-2001 (first entry)

XX Mycobacterium terrae rpoB gene fragment.

DE Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;

KW PCR-restriction fragment length polymorphism analysis; ds.

XX Mycobacterium terrae.

OS WO200131061-A1.

PN 03-MAY-2001.

XX 27-OCT-2000; 2000WO-KR01223.

XX 27-OCT-1999; 99KR-0046795.

PR (ERUM-) ERUME BIOTECH CO LTD.

PA Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;

XX WPI; 2001-300520/31.

XX New DNA fragments from the rpoB gene of mycobacteria, useful for

PT diagnosis and identification of many mycobacterial species by

PT restriction fragment length polymorphism -

PS Claim 1; Page 42; 50pp; English.

XX The present sequence for Mycobacterium terrae rpoB gene

CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from

CC various Mycobacterial species. These rpoB gene fragments can be used

CC in the diagnosis and identification of Mycobacterium species using a



novel PCR-restriction fragment length polymorphism analysis (PRA) method. The method comprises obtaining a restriction fragment length polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating, amplifying and digesting the DNA fragment from the microorganism to be identified and comparing the RFLP patterns from the known rpoB gene fragments with the unidentified fragment. The rpoB gene fragments are useful to identify a wide range of Mycobacterium species, e.g. for diagnosis or to obtain epidemiological and pathogenesis information for selection of appropriate therapies, including M. tuberculosis, M. leprae and non-tuberculous mycobacteria (NTM) encountered in subjects infected with human immunodeficiency virus (HIV). Analysis of the rpoB gene fragments is rapid, precise, simple and cost effective (only 1 PCR required), and can differentiate between many species in a single experiment, including those difficult to distinguish by usual biochemical tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for detecting specific Mycobacterial species.

Sequence 208 BP; 46 A; 62 C; 67 G; 33 T; 0 other;

Query Match 80.0%; Score 166.4; DB 22; Length 208;  
Best Local Similarity 87.5%; Pred. No. 6.5e-27;  
Matches 182; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 1 TCAAGGAGAGCGCTACGACCTGCGCCGGGTAGCCGCTACAAAGGTCAACAAGAGCTCG 60  
DB 1 TCAAGGAGAGCGCTACGACCTGCGCCGGGTAGCCGCTACAAAGGTCAACAAGAGCTCG 60  
QY 61 GCCTGCAGCTCGGGGATCCGATCACCAGCTCCACGCTGACCGAGGAGAGCTCGTCGCCA 120  
DB 61 GCCTGCATGTCGGCGAGCCCATACGTGCTGCAGCTGACCGAGGAGAGCTCGTCGCCA 120  
QY 121 CCATCGAGTACCTGTCGGCTGCACGAGGGCCAGCACAGCATGACCGTCCCGGGCGGCA 180  
DB 121 CCATCGAATATCTGTCGGCTGCACGAGGGTCAGACAGCATGACCGTTCGCGGGCGG 180  
QY 181 CCGAGGTGCGGGTTCAGACCGAGCAT 208  
DB 181 TCGAGGTGCGGGTGAACCGAGCAT 208

RESULT 13

AAS05216  
ID AAS05216 standard; DNA; 208 BP.

XX AAS05216;

DT 07-SEP-2001 (first entry)

DE Mycobacterium bovis rpoB gene fragment.

XX Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;  
KW PCR-restriction fragment length polymorphism analysis; ds.

XX Mycobacterium bovis.

XX WO200131061-A1.

XX 03-MAY-2001.

XX 27-OCT-2000; 2000WO-KR01223.

XX 27-OCT-1999; 99KR-0046795.

XX (ERUM-) ERUME BIOTECH CO LTD.

XX Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;

XX WPI; 2001-300520/31.

XX New DNA fragments from the rpoB gene of mycobacteria, useful for  
PT diagnosis and identification of many mycobacterial species by  
PT restriction fragment length polymorphism

PS Claim 1; Page 45; 50pp; English.

XX The present sequence for Mycobacterium bovis rpoB gene fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from various Mycobacterial species. These rpoB gene fragments can be used in the diagnosis and identification of Mycobacterium species using a novel PCR-restriction fragment length polymorphism analysis (PRA) method. The method comprises obtaining a restriction fragment length polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating, amplifying and digesting the DNA fragment from the microorganism to be identified and comparing the RFLP patterns from the known rpoB gene fragments with the unidentified fragment. The rpoB gene fragments are useful to identify a wide range of Mycobacterium species, e.g. for diagnosis or to obtain epidemiological and pathogenesis information for selection of appropriate therapies, including M. tuberculosis, M. leprae and non-tuberculous mycobacteria (NTM) encountered in subjects infected with human immunodeficiency virus (HIV). Analysis of the rpoB gene fragments is rapid, precise, simple and cost effective (only 1 PCR required), and can differentiate between many species in a single experiment, including those difficult to distinguish by usual biochemical tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for detecting specific Mycobacterial species.

XX Sequence 208 BP; 46 A; 62 C; 67 G; 33 T; 0 other;

Query Match 80.0%; Score 166.4; DB 22; Length 208;  
Best Local Similarity 87.5%; Pred. No. 6.5e-27;  
Matches 182; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 1 TCAAGGAGAGCGCTACGACCTGCGCCGGGTAGCCGCTACAAAGGTCAACAAGAGCTCG 60

DB 1 TCAAGGAGAGCGCTACGACCTGCGCCGGGTAGCCGCTACAAAGGTCAACAAGAGCTCG 60

QY 61 GCCTGCAGCTCGGGGATCCGATCACCAGCTCCACGCTGACCGAGGAGAGCTCGTCGCCA 120

DB 61 GCCTGCATGTCGGCGAGCCCATACGTGCTGCAGCTGACCGAGGAGAGCTCGTCGCCA 120

QY 121 CCATCGAGTACCTGTCGGCTGCACGAGGGCCAGCACAGCATGACCGTCCCGGGCGGCA 180

DB 121 CCATCGAATATCTGTCGGCTGCACGAGGGTCAGACAGCATGACCGTTCGCGGGCGG 180

QY 181 CCGAGGTGCGGGTTCAGACCGAGCAT 208

DB 181 TCGAGGTGCGGGTGAACCGAGCAT 208

RESULT 14

AAS05217  
ID AAS05217 standard; DNA; 208 BP.

XX AAS05217;

DT 07-SEP-2001 (first entry)

DE Mycobacterium celatum rpoB gene fragment.

XX Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;  
KW PCR-restriction fragment length polymorphism analysis; ds.

XX Mycobacterium celatum.

XX WO200131061-A1.

XX 03-MAY-2001.

XX 27-OCT-2000; 2000WO-KR01223.

XX 27-OCT-1999; 99KR-0046795.

XX (ERUM-) ERUME BIOTECH CO LTD.

XX Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;





PT with a primer set that targets portions of the gene encoding rpoB.

PS Disclosure; Fig.3; 54pp; English.

XX This oligonucleotide DNA primer is specific for Mycobacterium  
CC tuberculosis, and may be used to amplify a sample DNA by targeting  
CC a portion of the gene encoding rpoB. The 1st several bases comprise a  
CC nonhybridizing tail consisting of filler bases followed by  
CC a restriction site incorporated to facilitate cloning using the  
CC amplicon at a later date, if desired. The remaining bases hybridize  
CC to bacterial rpoB DNA. The method provides for the detection of M.  
CC tuberculosis and the concurrent determination of its drug  
CC susceptibility, particularly to rifamycin. The method can provide  
CC often greater than 95% sensitivity and 100% specificity. The  
CC biological sample is a fluid or tissue sample from a human.

XX Sequence 970 BP; 182 A; 302 C; 330 G; 156 T; 0 other;

Query Match 80.0%; Score 166.4; DB 17; Length 970;  
Best Local Similarity 87.5%; Pred. No. 6.3e-27;  
Matches 182; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

Qy	1	TCAAGGAGAACGCTACGACCTGCGCGGTAGGCGCTACAGGTCAACAAGAGCTCG	60
Db	26	TCAAGGAGAACGCTACGACCTGCGCGGTAGGCGCTACAGGTCAACAAGAGCTCG	85
Qy	61	GCCTGCAGTCGCGCGATCCGATCACCGCTCCAGCTGACCGAGGAGAGCTGCTGCCA	120
Db	86	GGCTGCATGTCGCGGAGCCCATACGTCGTCGACGCTGACCGAAGAGAGCTGCTGCCA	145
Qy	121	CCATCGAGTACCTGTCGCGCTGACGAGGCCAGCACAGCATGACCGTCCCGGGCGGCA	180
Db	146	CCATCGAATATCTGGTCCGCTTGACGAGGTCAGACCAGCATGACCGTTCGCGGGCGG	205
Qy	181	CCGAGGTCCCGGTTGAGACCGACAT	208
Db	206	TCGAGGTCCCGTGGAAACCGACAT	233

Search completed: November 12, 2002, 16:49:59  
Job time : 142.222 secs

GenCore Version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 12, 2002, 16:50:46 ; Search time 636.023 Seconds  
(without alignments)  
9517.553 Million cell updates/sec

Title: US-09-697-123B-1  
Perfect score: 208  
Sequence: 1 tcaaggagagcgctacgac.....ccggttgagaccgacgacat 208

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl:\*

- 1: gb\_ba:\*
- 2: gb\_htg:\*
- 3: gb\_in:\*
- 4: gb\_om:\*
- 5: gb\_ov:\*
- 6: gb\_pat:\*
- 7: gb\_ph:\*
- 8: gb\_pl:\*
- 9: gb\_pr:\*
- 10: gb\_ro:\*
- 11: gb\_sts:\*
- 12: gb\_sy:\*
- 13: gb\_un:\*
- 14: gb\_vl:\*
- 15: em\_ba:\*
- 16: em\_fun:\*
- 17: em\_hum:\*
- 18: em\_in:\*
- 19: em\_mu:\*
- 20: em\_om:\*
- 21: em\_or:\*
- 22: em\_ov:\*
- 23: em\_pat:\*
- 24: em\_ph:\*
- 25: em\_pl:\*
- 26: em\_ro:\*
- 27: em\_sts:\*
- 28: em\_un:\*
- 29: em\_vl:\*
- 30: em\_htg\_hum:\*
- 31: em\_htg\_inv:\*
- 32: em\_htg\_other:\*
- 33: em\_htg\_mus:\*
- 34: em\_htg\_pln:\*
- 35: em\_htg\_rod:\*
- 36: em\_htg\_mam:\*
- 37: em\_htg\_vrt:\*
- 38: em\_sv:\*
- 39: em\_htgo\_hum:\*
- 40: em\_htgo\_mus:\*
- 41: em\_htgo\_other:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	166.4	80.0	610	1	MTU318818	AJ318818 Mycobacte
2	166.4	80.0	610	1	MTU318819	AJ318819 Mycobacte
3	166.4	80.0	616	1	MTU318813	AJ318813 Mycobacte
4	166.4	80.0	618	1	MTU318815	AJ318815 Mycobacte
5	166.4	80.0	618	1	MTU318817	AJ318817 Mycobacte
6	166.4	80.0	633	1	MTU318814	AJ318814 Mycobacte
7	166.4	80.0	637	1	MTU318816	AJ318816 Mycobacte
8	166.4	80.0	639	1	MTU318821	AJ318821 Mycobacte
9	166.4	80.0	970	6	IS0706	Sequence 1
10	166.4	80.0	3534	6	AX111339	Sequence
11	166.4	80.0	3853	1	MTU12205	U12205 Mycobacteri
12	166.4	80.0	5084	1	MSGRPOB	L27989 Mycobacteri
13	166.4	80.0	19352	1	AE006964	AE006964 Mycobacte
14	166.4	80.0	19770	1	MTCI376	295972 Mycobacteri
15	160	76.9	3752	1	MSU24494	U24494 Mycobacteri
16	142.4	68.5	3447	6	AR067447	AR067447 Sequence
17	142.4	68.5	37617	1	MLB1790G	214314 M.leprae ge
18	142.4	68.5	37617	1	MLB1790G	AL583923 Mycobacte
19	122	58.7	3941	1	AF242549	AF242549 Amycolato
20	91.6	44.0	32923	1	SCD82	AL160431 Streptomy
21	83.2	40.0	3495	6	AX120631	AX120631 Sequence
22	83.2	40.0	328050	1	AP005275	AP005275 Corynebac
23	83.2	40.0	349980	6	AX127144	AX127144 Sequence 4
24	49.8	23.9	6085	6	A60304	A60304 Sequence 4
25	49.8	23.9	6085	6	ARI44763	ARI44763 Sequence
26	46.2	22.2	12112	1	AE005722	AE005722 Caulobact
27	46.2	22.2	347660	1	AP002994	AP002994 Mesorhizo
28	45.8	22.0	10846	1	AE005837	AE005837 Caulobact
29	45.8	22.0	329100	1	SME591787	AL591787 Sinorhizo
30	45.2	21.7	2100	9	AK096697	AK096697 Homo sapi
31	45.2	21.7	118312	9	AC055713	AC055713 Homo sapi
32	45	21.6	3137	8	AF268091	AF268091 Chloris g
33	44.2	21.3	7954	1	TAQ19223	Y19223 Thermus aqu
34	44.2	21.3	12523	1	AE001898	AE001898 Deinococc
35	44.2	21.3	16160	1	SC6C5	AL034492 Streptomy
36	44.2	21.3	20054	1	SCCB12	AL591588 Streptomy
37	44.2	21.3	134816	2	CNS08C9R	AL772413 Oryza sat
38	44.2	21.3	135378	2	CNS07XP9	AL713941 Oryza sat
39	44.2	21.3	196050	1	AL646058	AL646058 Ralstonia
40	44	21.2	299350	1	SME591786	AL591786 Sinorhizo
41	43.8	21.1	2966	8	AF247649	AF247649 Coniothyr
42	43.8	21.1	13818	1	AE004747	AE004747 Pseudomon
43	43.8	21.1	33095	1	SC3C8	AL023861 Streptomy
44	43.6	21.0	127108	8	OSJN00144	AL662949 Oryza sat
45	43.6	21.0	130014	2	OSJN00214	AL663013 Oryza sat

ALIGNMENTS

RESULT 1

MTU318818

LOCUS

DEFINITION

MTU318818

610 bp DNA circular BCT 09-AUG-2002

Myco

tuberculosis partial rpoB gene for RNA polymerase

beta subunit, isolate 1415-97.

ACCESSION

AJ318818

VERSION

AJ318818.1

GI:22208412

KEYWORDS

RNA polymerase beta subunit; rpoB gene.

SOURCE

Myco

tuberculosis

ORGANISM

Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;

Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium

tuberculosis complex.

REFERENCE

1

Herrera, L., Jimenez, M.S. and Saez, J.A.

TITLE Molecular analysis of rifampin-resistant Mycobacterium tuberculosis isolated in Spain (1996-2001). Description of new alleles into rpoB gene and review

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 610)

AUTHORS Herrera, L.

TITLE Direct Submission

JOURNAL Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5, Majadahonda. Madrid. 28220, SPAIN

FEATURES

source

1. .610

/organism="Mycobacterium tuberculosis"

/isolate="1415-97"

/db\_xref="taxon:1773"

gene

1. .610

/gene="rpoB"

CDS

<1. .>610

/codon\_start=1

/transl\_table=11

/product="RNA polymerase beta subunit"

/protein\_id="CAC87035.1"

/db\_xref="GI:22208413"

/translation="LDIYRKLPGPEPTKESAOITLLENLFFKEKRYDLARVGRYKVNK KLGLHVGEPIITSLTIEDVATIEYLRHGGQTTHTPGGEVPEVETDDIDHFGNR RLRTVGEIIONQIRVGMSRMERVMRTQDVEAITPQTILINIRPVAAIKEFFGTS QPSQFMQNPLSLGTHKRRLSALGPGGLSRERAGLEVRDHP"

BASE COUNT 122 a 191 c 202 g 95 t

ORIGIN

Query Match 80.0%; Score 166.4; DB 1; Length 610;

Best Local Similarity 87.5%; Pred. No. 1.5e-16;

Matches 182; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 1 TCAAGAGAAGCGCTACGACCTGGCCGGGTAGCGCGCTACAAAGTCAACAGAAGCTCG 60

DB 80 TCAAGAGAAGCGCTACGACCTGGCCGGGTAGCGCGCTACAAAGTCAACAGAAGCTCG 139

QY 61 GCCTGCACCTCGCGATCCAGCTCCAGCTGACCGGAGGAGAGAGCTCGTGCCTCA 120

DB 140 GGTCTGATGTCGGCGAGCCCATCGCTGCTGCTATTAAGTCAACAGAAGCTCG 199

QY 121 CCATCGAGTACCTGTCGCCCTCGACGAGGCGCAGCACAGATGACCGTCCCGGGCGGCA 180

DB 260 TCGAGTGCCGGTGAACCGAGACAT 287

QY 181 CCGAGTGCCGGTGAACCGAGACAT 208

DB 260 TCGAGTGCCGGTGAACCGAGACAT 287

RESULT 2

MTU318819

LOCUS

DEFINITION

610 bp DNA circular BCT 09-AUG-2002

Myco bacterium tuberculosis partial rpoB gene for RNA polymerase beta subunit, isolate 1417-97.

ACCESSION

AJ318819

VERSION

AJ318819.1 GI:22208414

KEYWORDS

RNA polymerase beta subunit; rpoB gene.

SOURCE

Myco bacterium tuberculosis.

ORGANISM

Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium tuberculosis complex.

REFERENCE

1

AUTHORS

Herrera, L., Jimenez, M.S. and Saez, J.A.

TITLE

Molecular analysis of rifampin-resistant Mycobacterium tuberculosis isolated in Spain (1996-2001). Description of new alleles into rpoB gene and review

JOURNAL

Unpublished

REFERENCE

2 (bases 1 to 610)

AUTHORS

Herrera, L.

TITLE Direct Submission

JOURNAL Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5, Majadahonda. Madrid. 28220, SPAIN

FEATURES

source

1. .610

/organism="Mycobacterium tuberculosis"

/isolate="1417-97"

/db\_xref="taxon:1773"

gene

1. .610

/gene="rpoB"

CDS

<1. .>610

/codon\_start=1

/transl\_table=11

/product="RNA polymerase beta subunit"

/protein\_id="CAC87036.1"

/db\_xref="GI:22208415"

/translation="LDIYRKLPGPEPTKESAOITLLENLFFKEKRYDLARVGRYKVNK KLGLHVGEPIITSLTIEDVATIEYLRHGGQTTHTPGGEVPEVETDDIDHFGNR RLRTVGEIIONQIRVGMSRMERVMRTQDVEAITPQTILINIRPVAAIKEFFGTS QPSQFMQNPLSLGTHKRRLSALGPGGLSRERAGLEVRDHP"

BASE COUNT 122 a 191 c 202 g 95 t

ORIGIN

Query Match 80.0%; Score 166.4; DB 1; Length 610;

Best Local Similarity 87.5%; Pred. No. 1.5e-16;

Matches 182; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 1 TCAAGAGAAGCGCTACGACCTGGCCGGGTAGCGCGCTACAAAGTCAACAGAAGCTCG 60

DB 80 TCAAGAGAAGCGCTACGACCTGGCCGGGTAGCGCGCTACAAAGTCAACAGAAGCTCG 139

QY 61 GCCTGCACCTCGCGATCCAGCTCCAGCTGACCGGAGGAGAGAGCTCGTGCCTCA 120

DB 140 GGTCTGATGTCGGCGAGCCCATCGCTGCTGCTATTAAGTCAACAGAAGCTCG 199

QY 121 CCATCGAGTACCTGTCGCCCTCGACGAGGCGCAGCACAGATGACCGTCCCGGGCGGCA 180

DB 200 CCATCGAATATCTGTCGCCCTGTCACGAGGCTCAGACCGATGACCGTTCGGGGCGG 259

QY 181 CCGAGTGCCGGTGAACCGAGACAT 208

DB 260 TCGAGTGCCGGTGAACCGAGACAT 287

RESULT 3

MTU318813

LOCUS

DEFINITION

616 bp DNA circular BCT 09-AUG-2002

Myco bacterium tuberculosis partial rpoB gene for RNA polymerase beta subunit, isolate 1763-97.

ACCESSION

AJ318813

VERSION

AJ318813.1 GI:22208402

KEYWORDS

RNA polymerase beta subunit; rpoB gene.

SOURCE

Myco bacterium tuberculosis.

ORGANISM

Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium tuberculosis complex.

REFERENCE

1

AUTHORS

Herrera, L., Jimenez, M.S. and Saez, J.A.

TITLE

Molecular analysis of rifampin-resistant Mycobacterium tuberculosis isolated in Spain (1996-2001). Description of new alleles into rpoB gene and review

JOURNAL

Unpublished

REFERENCE

2 (bases 1 to 616)

AUTHORS

Herrera, L.

TITLE

Direct Submission

JOURNAL

Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5, Majadahonda. Madrid. 28220, SPAIN

FEATURES

source

1. .616

gene /organism="Mycobacterium tuberculosis"  
/isolate="1763-97"  
/db\_xref="taxon:1773"

CDS 1..615  
/gene="rpoB"  
/codon\_start=1  
/transl\_table=11  
/product="RNA polymerase beta subunit"  
/protein\_id="CAC87030.1"  
/db\_xref="GI:22208403"

BASE COUNT 125 a 191 c 201 g 99 t  
ORIGIN  
Query Match 80.0%; Score 166.4; DB 1; Length 616;  
Best Local Similarity 87.5%; Pred. No. 1.5e-16;  
Matches 182; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

Qy 1 TCAAGGAGAGCGCTACGACCTGGCCCGGTAGCGCTACAAAGTCAACAAAGAGCTCG 60  
Db 80 TCAAGGAGAGCGCTACGACCTGGCCCGGTAGCGCTACAAAGTCAACAAAGAGCTCG 139

Qy 61 GCCTGCACGTCGGGATCCGATCACCAGCTCCACGCTGACCGAGGAGAGCTCGTCGCCA 120  
Db 140 GCCTGCATGTCGGGAGCGCCATCAGCTGTCGACGCTGACCGAGGAGAGCTCGTCGCCA 199

Qy 121 CCATCGAGTACCTGGTCCGCTCCAGGAGGCGCACACGATGACCGTCCCGGGCGGCA 180  
Db 200 CCATCGAATATCTGCTCGCTCCAGGAGGTCAGACCATGACCGTCCCGGGCGGCG 259

Qy 181 CCGAGGTGCGGTTGAGACCGACGACAT 208  
Db 260 TCGAGGTGCGGTTGGAACCGACGACAT 287

RESULT 4  
MTU318815 618 bp DNA circular BCT 09-AUG-2002  
LOCUS Mycobacterium tuberculosis partial rpoB gene for RNA polymerase  
DEFINITION beta subunit, isolate 2540-97.

ACCESSION AJ318815  
VERSION AJ318815.1 GI:22208406  
KEYWORDS RNA polymerase beta subunit; rpoB gene.  
SOURCE Mycobacterium tuberculosis.  
ORGANISM Bacteria; Actinobacteria; Actinobacteridae; Mycobacterium; Corynebacterineae; Mycobacteriaceae; Mycobacterium tuberculosis complex.

REFERENCE 1  
Herrera,L., Jimenez,M.S. and Saez,J.A.  
AUTHORS Molecular analysis of rifampin-resistant Mycobacterium tuberculosis  
TITLE Isolated in Spain (1996-2001). Description of new alleles into rpoB  
JOURNAL gene and review

REFERENCE 2 (bases 1 to 618)  
Herrera,L.  
AUTHORS Direct Submission  
TITLE Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro  
JOURNAL Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5,  
Majadahonda. Madrid. 28220, SPAIN

FEATURES  
source 1..618  
/organism="Mycobacterium tuberculosis"  
/isolate="2540-97"  
/db\_xref="taxon:1773"

gene 1..618  
/gene="rpoB"  
CDS <1..>618

/gene="rpoB"  
/codon\_start=1  
/transl\_table=11  
/product="RNA polymerase beta subunit"  
/protein\_id="CAC87032.1"  
/db\_xref="GI:22208407"

BASE COUNT 124 a 190 c 207 g 97 t  
ORIGIN  
Query Match 80.0%; Score 166.4; DB 1; Length 618;  
Best Local Similarity 87.5%; Pred. No. 1.5e-16;  
Matches 182; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

Qy 1 TCAAGGAGAGCGCTACGACCTGGCCCGGTAGCGCTACAAAGTCAACAAAGAGCTCG 60  
Db 95 TCAAGGAGAGCGCTACGACCTGGCCCGGTAGCGCTACAAAGTCAACAAAGAGCTCG 154

Qy 61 GCCTGCACGTCGGGATCCGATCACCAGCTCCACGCTGACCGAGGAGAGCTCGTCGCCA 120  
Db 155 GCCTGCATGTCGGGAGCGCCATCAGCTGTCGACGCTGACCGAGGAGAGCTCGTCGCCA 214

Qy 121 CCATCGAGTACCTGGTCCGCTCCAGGAGGCGCACACGATGACCGTCCCGGGCGGCA 180  
Db 215 CCATCGAATATCTGCTCGCTTCAGGAGGTCAGACCATGACCGTCCCGGGCGGCG 274

Qy 181 CCGAGGTGCGGTTGAGACCGACGACAT 208  
Db 275 TCGAGGTGCGGTTGGAACCGACGACAT 302

RESULT 5  
MTU318817 618 bp DNA circular BCT 09-AUG-2002  
LOCUS Mycobacterium tuberculosis partial rpoB gene for RNA polymerase  
DEFINITION beta subunit, isolate 2348-98.

ACCESSION AJ318817  
VERSION AJ318817.1 GI:22208410  
KEYWORDS RNA polymerase beta subunit; rpoB gene.  
SOURCE Mycobacterium tuberculosis.  
ORGANISM Mycobacterium tuberculosis  
Bacteria; Actinobacteria; Actinobacteridae; Mycobacterium; Corynebacterineae; Mycobacteriaceae; Mycobacterium tuberculosis complex.

REFERENCE 1  
Herrera,L., Jimenez,M.S. and Saez,J.A.  
AUTHORS Molecular analysis of rifampin-resistant Mycobacterium tuberculosis  
TITLE Isolated in Spain (1996-2001). Description of new alleles into rpoB  
JOURNAL gene and review

REFERENCE 2 (bases 1 to 618)  
Herrera,L.  
AUTHORS Direct Submission  
TITLE Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro  
JOURNAL Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5,  
Majadahonda. Madrid. 28220, SPAIN

FEATURES  
source 1..618  
/organism="Mycobacterium tuberculosis"  
/isolate="2348-98"  
/db\_xref="taxon:1773"

gene 1..618  
/gene="rpoB"  
CDS <1..>618

/gene="rpoB"  
/codon\_start=1  
/transl\_table=11  
/product="RNA polymerase beta subunit"  
/protein\_id="CAC87034.1"  
/db\_xref="GI:22208411"

/translation="TDEALLDIYKLRPCGPPTKESAOITLLENLFKEKRYDLRVGR  
YKVNKKLGLHVGEPTTSSTLTEDVVAITYEYLVRLHEGQTTMTVPGGVEPVETDDID  
HFGNRLRTVGLIQIRVGMRSRMVVRMTTQDVEAITPOTLINIRPVVAIKE  
FFGTSPQFMGONNPLSLGTHKRLSALPGGLSRERAGLEVRDV"  
BASE COUNT 124 a 192 c 207 g 95 t  
ORIGIN

Query Match 80.0%; Score 166.4; DB 1; Length 618;  
Best Local Similarity 87.5%; Pred. No. 1.5e-16;  
Matches 182; Conservative 0; Mismatches 26; Indels 0; Gaps 0;  
QY 1 TCAAGGAGAGCGCTACGACCTGCGCGGTAGCGCTACAAGGTCAACAAGAACTCG 60  
DB 95 TCAAGGAGAGCGCTACGACCTGCGCGGTAGCGCTACAAGGTCAACAAGAACTCG 154  
QY 61 GCCTGCAGTTCGGGATCCCATCACCAGTTCACGCTGACCGAGGAGAGAGCTCGCGCCA 120  
DB 155 GGCTGCATGTGCGGAGCGCCATCAGCTGCTGACCGTGCACGAGGAGAGAGCTCGTGCCCA 214  
QY 121 CCATCGAGTACCTGCTGCGCTGCACGAGGCGCCAGCACAGCATGACCGTCCCGGGCGGCA 180  
DB 215 CCATCGAATATCTGCTCGCTGCACGAGGCTCAGCAGCATGACCGTTCGGGGCGGCG 274  
QY 181 CCGAGTGCCTGCTGAGACCGACGACAT 208  
DB 275 TCGAGTGCCTGCTGAGAACCGACGACAT 302

RESULT 6  
MTU318814  
LOCUS  
DEFINITION  
MTU318814  
633 bp DNA circular BCT 09-AUG-2002  
Mycobacterium tuberculosis partial rpoB gene for RNA polymerase  
beta subunit, isolate 1058-97.  
ACCESSION  
AJ318814  
VERSION  
AJ318814.1 GI:22208404  
KEYWORDS  
RNA polymerase beta subunit; rpoB gene.  
SOURCE  
Mycobacterium tuberculosis.  
ORGANISM  
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium  
tuberculosis complex.

REFERENCE  
1  
Herrera, L., Jimenez, M.S. and Saez, J.A.  
Molecular analysis of rifampin-resistant Mycobacterium tuberculosis  
isolated in Spain (1996-2001). Description of new alleles into rpoB  
gene and review  
Unpublished  
2 (bases 1 to 633)  
Herrera, L.  
Direct Submission  
Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro  
Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5,  
Majadahonda, Madrid. 28220, SPAIN

FEATURES  
source  
1..633  
/organism="Mycobacterium tuberculosis"  
/isolate="1058-97"  
/db\_xref="taxon:1773"  
gene  
1..633  
/gene="rpoB"  
CDS  
1..>633  
/gene="rpoB"  
/codon\_start=1  
/translation="DNTVGTDEALLDIYKLRPCGPPTKESAOITLLENLFKEKRYDL  
ARVGRYKVNKKLGLHVGEPTTSSTLTEDVVAITYEYLVRLHEGQTTMTVPGGVEPV  
ETDDIDHFGNRLRTVGLIQIRVGMRSRMVVRMTTQDVEAITPOTLINIRPVV  
AAIKEFFTSQFSQFMGONNPLSLGTHKRLSALPGGLSRERAGLEVRDV"  
BASE COUNT 129 a 195 c 210 g 99 t  
ORIGIN

Query Match 80.0%; Score 166.4; DB 1; Length 633;  
Best Local Similarity 87.5%; Pred. No. 1.5e-16;  
Matches 182; Conservative 0; Mismatches 26; Indels 0; Gaps 0;  
QY 1 TCAAGGAGAGCGCTACGACCTGCGCGGTAGCGCTACAAGGTCAACAAGAACTCG 60  
DB 110 TCAAGGAGAGCGCTACGACCTGCGCGGTAGCGCTACAAGGTCAACAAGAACTCG 169  
QY 61 GCCTGCAGTTCGGGATCCCATCACCAGTTCACGCTGACCGAGGAGAGAGCTCGTGCCCA 120  
DB 170 GGCTGCATGTGCGGAGCGCCATCAGCTGCTGACCGTGCACGAGGAGAGAGCTCGTGCCCA 229  
QY 121 CCATCGAGTACCTGCTGCGCTGCACGAGGCGCCAGCACAGCATGACCGTCCCGGGCGGCA 180  
DB 230 CCATCGAATATCTGCTGCGCTGCACGAGGCTCAGACCAACGATGACCGTTCGGGGCGGCG 289  
QY 181 CCGAGTGCCTGCTGAGACCGACGACAT 208  
DB 290 TCGAGTGCCTGCTGAGAACCGACGACAT 317

RESULT 7  
MTU318816  
LOCUS  
DEFINITION  
MTU318816  
637 bp DNA circular BCT 09-AUG-2002  
Mycobacterium tuberculosis partial rpoB gene for RNA polymerase  
beta subunit, isolate 1255-98.  
ACCESSION  
AJ318816  
VERSION  
AJ318816.1 GI:22208408  
KEYWORDS  
RNA polymerase beta subunit; rpoB gene.  
SOURCE  
Mycobacterium tuberculosis.  
ORGANISM  
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium  
tuberculosis complex.

REFERENCE  
1  
Herrera, L., Jimenez, M.S. and Saez, J.A.  
Molecular analysis of rifampin-resistant Mycobacterium tuberculosis  
isolated in Spain (1996-2001). Description of new alleles into rpoB  
gene and review  
Unpublished  
2 (bases 1 to 637)  
Herrera, L.  
Direct Submission  
Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro  
Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5,  
Majadahonda, Madrid. 28220, SPAIN

FEATURES  
source  
1..637  
/organism="Mycobacterium tuberculosis"  
/isolate="1255-98"  
/db\_xref="taxon:1773"  
gene  
1..637  
/gene="rpoB"  
CDS  
1..>637  
/gene="rpoB"  
/codon\_start=1  
/transl\_table=11  
/product="RNA polymerase beta subunit"  
/protein\_id="CAC87033.1"  
/db\_xref="GI:22208409"  
/translation="VGTDEALLDIYKLRPCGPPTKESAOITLLENLFKEKRYDLARV  
GRYKVNKKLGLHVGEPTTSSTLTEDVVAITYEYLVRLHEGQTTMTVPGGVEPVETDD  
IDHFGNRLRTVGLIQIRVGMRSRMVVRMTTQDVEAITPOTLINIRPVVA  
AIIKEFFTSQFSQFMGONNPLSLGTHKRLSALPGGLSRERAGLEVRDVHPSH"  
BASE COUNT 128 a 198 c 210 g 101 t  
ORIGIN

Query Match 80.0%; Score 166.4; DB 1; Length 637;  
Best Local Similarity 87.5%; Pred. No. 1.5e-16;  
Matches 182; Conservative 0; Mismatches 26; Indels 0; Gaps 0;  
QY 1 TCAAGGAGAGCGCTACGACCTGCGCGGTAGCGCTACAAGGTCAACAAGAACTCG 60

|||||  
Db 101 TCAAGGAGAGCGCTACGACCTGGCCCGCTCGGTGCTATTAAGTCAACAGAGAGCTCG 160  
QY 61 GCCTGCAGCTGCGGATCCGATCACCAGCTCCACGCTGACCGAGGAGACGCTGTCGGCCA 120  
Db 161 GGCTGCATGTCGGCGAGCGCATCACCTCGCTCGACGCTACCGAAGAGAGCTGTCGGCCA 220  
QY 121 CCATCGAGTACCTGTCGCGCTGCACGAGGCGCACACAGCATGACCGTCCCGGGCGGCA 180  
Db 221 CCATCGAATATCTGGTCCGCTTGCACGAGGGTGCACACGATGACCGTTCGGGGCGGCG 280  
QY 181 CCGAGTGCCTGGTTGAGACCGACGACAT 208  
Db 281 TCGAGTGCCTGGTGAACCGAGGACAT 308  
RESULT 8  
MTU318821 639 bp DNA circular BCT 09-AUG-2002  
LOCUS Mycobacterium tuberculosis partial rpoB gene for RNA polymerase  
DEFINITION beta subunit, isolate 1071-98.  
ACCESSION AJ318821  
VERSION AJ318821.1 GI:22208418  
KEYWORDS RNA polymerase beta subunit; rpoB gene.  
SOURCE Mycobacterium tuberculosis.  
ORGANISM Mycobacterium tuberculosis.  
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium  
tuberculosis complex.  
REFERENCE 1  
AUTHORS Herrera, L., Jimenez, M.S. and Saez, J.A.  
TITLE Molecular analysis of rifampin-resistant Mycobacterium tuberculosis  
isolated in Spain (1996-2001). Description of new alleles into rpoB  
gene and review  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 639)  
AUTHORS Herrera, L.  
TITLE Direct Submission  
JOURNAL Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro  
Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5,  
Majadahonda. Madrid. 28220, SPAIN  
FEATURES  
Source  
1. 639  
/organism="Mycobacterium tuberculosis"  
/isolate="1071-98"  
/db\_xref="taxon:1773"  
gene 1. .639  
/gene="rpoB"  
CDS <1..>639  
/gene="rpoB"  
/codon\_start=1  
/transl\_table=1  
/product="RNA polymerase beta subunit"  
/protein\_id="CAC87038.1"  
/db\_xref="GI:22208419"  
/translation="ARTDEALLDIYRKLPGCEPTKESQTLLENLFKRYDLARV  
IDYKNKGLHVGPISTLTEDVATIEYLVRHGGTTMTVPGVSPVETDD  
IHFQGNRLRTVGLIQNOIRVMSRMYRRTQDVEAITPQTLINRPAAAI  
KEFFTQSLSQFMVQNNPLSLGTQKRLSALPGLSRERAGLEVDRVHPSHS"  
BASE COUNT 126 a 202 c 212 g 99 t  
ORIGIN  
Query Match 80.0%; Score 166.4; DB 1; Length 639;  
Best Local Similarity 87.5%; Pred. No. 1.5e-16;  
Matches 182; Conservative 0; Mismatches 26; Indels 0; Gaps 0;  
QY 1 TCAAGGAGAGCGCTACGACCTGGCCCGGTAGCGGCTACAGGTCAACAGAGAGCTCG 60  
Db 101 TCAAGGAGAGCGCTACGACCTGGCCCGGTAGCGGCTATTAAGTCAACAGAGAGCTCG 160  
QY 61 GCCTGCAGCTGCGGATCCGATCACCAGCTCCACGCTGACCGAGGAGACGCTGTCGGCCA 120  
Db 161 GGCTGCATGTCGGCGAGCGCATCACCTCGCTCGACGCTACCGAAGAGAGCTGTCGGCCA 220  
QY 121 CCATCGAGTACCTGTCGCGCTGCACGAGGCGCACACAGCATGACCGTCCCGGGCGGCA 180  
Db 221 CCATCGAATATCTGGTCCGCTTGCACGAGGGTGCACACGATGACCGTTCGGGGCGGCG 280  
QY 181 CCGAGTGCCTGGTTGAGACCGACGACAT 208  
Db 281 TCGAGTGCCTGGTGAACCGAGGACAT 308  
RESULT 8  
MTU318821 639 bp DNA circular BCT 09-AUG-2002  
LOCUS Mycobacterium tuberculosis partial rpoB gene for RNA polymerase  
DEFINITION beta subunit, isolate 1071-98.  
ACCESSION AJ318821  
VERSION AJ318821.1 GI:22208418  
KEYWORDS RNA polymerase beta subunit; rpoB gene.  
SOURCE Mycobacterium tuberculosis.  
ORGANISM Mycobacterium tuberculosis.  
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium  
tuberculosis complex.  
REFERENCE 1  
AUTHORS Herrera, L., Jimenez, M.S. and Saez, J.A.  
TITLE Molecular analysis of rifampin-resistant Mycobacterium tuberculosis  
isolated in Spain (1996-2001). Description of new alleles into rpoB  
gene and review  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 639)  
AUTHORS Herrera, L.  
TITLE Direct Submission  
JOURNAL Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro  
Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5,  
Majadahonda. Madrid. 28220, SPAIN  
FEATURES  
Source  
1. 639  
/organism="Mycobacterium tuberculosis"  
/isolate="1071-98"  
/db\_xref="taxon:1773"  
gene 1. .639  
/gene="rpoB"  
CDS <1..>639  
/gene="rpoB"  
/codon\_start=1  
/transl\_table=1  
/product="RNA polymerase beta subunit"  
/protein\_id="CAC87038.1"  
/db\_xref="GI:22208419"  
/translation="ARTDEALLDIYRKLPGCEPTKESQTLLENLFKRYDLARV  
IDYKNKGLHVGPISTLTEDVATIEYLVRHGGTTMTVPGVSPVETDD  
IHFQGNRLRTVGLIQNOIRVMSRMYRRTQDVEAITPQTLINRPAAAI  
KEFFTQSLSQFMVQNNPLSLGTQKRLSALPGLSRERAGLEVDRVHPSHS"  
BASE COUNT 126 a 202 c 212 g 99 t  
ORIGIN  
Query Match 80.0%; Score 166.4; DB 1; Length 639;  
Best Local Similarity 87.5%; Pred. No. 1.5e-16;  
Matches 182; Conservative 0; Mismatches 26; Indels 0; Gaps 0;  
QY 1 TCAAGGAGAGCGCTACGACCTGGCCCGGTAGCGGCTACAGGTCAACAGAGAGCTCG 60  
Db 101 TCAAGGAGAGCGCTACGACCTGGCCCGGTAGCGGCTATTAAGTCAACAGAGAGCTCG 160  
QY 61 GCCTGCAGCTGCGGATCCGATCACCAGCTCCACGCTGACCGAGGAGACGCTGTCGGCCA 120  
Db 161 GGCTGCATGTCGGCGAGCGCATCACCTCGCTCGACGCTACCGAAGAGAGCTGTCGGCCA 220

QY 121 CCATCGAGTACCTGTCGCTCGCTGCACGAGGCGCACACAGCATGACCGTCCCGGGCGGCA 180  
Db 221 CCATCGAATATCTGTCGCTTGCACGAGGGTGCAGACGATGACCGTTCGGGGCGGCG 280  
QY 181 CCGAGGTCGCGGTTGAGACCGACGACAT 208  
Db 281 TCGAGGTCGCGGTGAAACCGACGACAT 308  
RESULT 9  
150706 970 bp DNA linear PAT 07-OCT-1997  
LOCUS Sequence 1 from patent US 5643723.  
DEFINITION 150706  
ACCESSION 150706  
VERSION 150706.1 GI:2472409  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 970)  
AUTHORS Persing, D.H., Hunt, J.J., Young, K.Y., Feimlee, T.A., Roberts, G.D.  
and Whelan, A.Christian.  
TITLE Detection of a genetic locus encoding resistance to rifampin in  
mycobacterial cultures and in clinical specimens  
JOURNAL Patent: US 5643723-A 1 01-JUL-1997;  
FEATURES Location/Qualifiers  
source 1. .970  
/organism="unknown"  
BASE COUNT 182 a 302 c 330 g 156 t  
ORIGIN  
Query Match 80.0%; Score 166.4; DB 6; Length 970;  
Best Local Similarity 87.5%; Pred. No. 1.4e-16;  
Matches 182; Conservative 0; Mismatches 26; Indels 0; Gaps 0;  
QY 1 TCAAGGAGAGCGCTAGACCTGCGCCGCTAGCGGCTAGCCGCTACAAGGTCAACAAGAGCTCG 60  
Db 26 TCAAGGAGAGCGCTAGACCTGCGCCGCTAGCGGCTAGCGGCTACAAGGTCAACAAGAGCTCG 85  
QY 61 GCCTGCAGCTCGCGGATCCGATCACCAGCTCCACGCTGACCGAGGAGAGAGCTGTCGCCCA 120  
Db 86 GCCTGCATGTCGGCGAGCGCCATCACGTCGTCGACGCTGACCGAAGAGAGCTGTCGCCCA 145  
QY 121 CCATCGAGTACCTGTCGCTCGCTGCACGAGGCGCACACAGCATGACCGTCCCGGGCGGCA 180  
Db 146 CCATCGAATATCTGTCGCTTGCACGAGGGTGCAGACGATGACCGTTCGGGGCGGCG 205  
QY 181 CCGAGGTCGCGGTTGAGACCGACGACAT 208  
Db 206 TCGAGGTCGCGGTGAAACCGACGACAT 233  
RESULT 10  
AX111339 3534 bp DNA linear PAT 30-APR-2001  
LOCUS Sequence 2072 from Patent WO0123604.  
DEFINITION AX111339  
ACCESSION AX111339  
VERSION AX111339.1 GI:13927631  
KEYWORDS  
SOURCE Mycobacterium tuberculosis.  
ORGANISM Mycobacterium tuberculosis.  
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
Actinomycetales; Corynebacterineae; Mycobacteriaceae;  
Mycobacterium; Mycobacterium tuberculosis complex.  
REFERENCE 1 (bases 1 to 3534)  
AUTHORS Bergeron, M.G., Boissinot, M., Hulet, A., m Nard, C., Ouellette, M.,  
Picard, F.J. and Roy, P.H.  
TITLE Highly conserved genes and their use to generate probes and primers  
for detection of microorganisms  
JOURNAL Patent: WO 0123604-A 2072 05-APR-2001;  
FEATURES Infectio Diagnostic (I.D.I.) INC. (CA)  
Location/Qualifiers

source	1. .3534	/organism="Mycobacterium tuberculosis"	
		/strain="Rv"	
BASE COUNT	679 a 1081 c 1188 g 586 t		
ORIGIN			
Query Match	80.0%;	Score 166.4;	DB 6; Length 3534;
Best Local Similarity	87.5%;	Pred. No. 1.1e-16;	
Matches 182;	Conservative 0;	Mismatches 26;	Indels 0; Gaps 0;
Qy 1	TCAAGGAGAAGCGCTAGCAGCTGCCCGGTAGCCGCTACAGGCTCAAGGTCACAAAGAAGCTCG	60	
Db 902	TCAAGGAGAAGCGCTAGCAGCTGCCCGGTAGCCGCTACAGGCTCAAGGTCACAAAGAAGCTCG	961	
Qy 61	GCCTGCAGCTGCGGATCCGATCACCAGCTCCAGCTCCAGGAGAGAGCTCGTCGCCCA	120	
Db 962	GCCTGCATGTCCGGAGGCCATCAGCTCGCTGCGCTGACCGAAGAGAGCTCGTCGCCCA	1021	
Qy 121	CCATCGAGTACCTCGCTCGCTGACGAGGCCAGCACGATCACCCTCCCGGGCGCA	180	
Db 1022	CCATCGAATATCTCGCTCGCTGACGAGGTCAGACCAGATGACCGTTCCGGGCGCG	1081	
Qy 181	CCGAGGTCCCGTTGAGACCGACGACAT	208	
Db 1082	TCGAGGTCCCGTGGAAACCGACGACAT	1109	
RESULT 11			
MTU12205			
LOCUS	3853 bp	DNA	linear BCT 02-MAR-2000
DEFINITION	Mycobacterium tuberculosis H37Rv RNA-polymerase beta subunit (rpoB) gene, partial cds.		
ACCESSION	U12205		
VERSION	U12205.1	GI:515684	
KEYWORDS			
SOURCE			
ORGANISM	Mycobacterium tuberculosis.		
	Mycobacterium tuberculosis		
	Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;		
	Actinomycetales; Corynebacterineae; Mycobacteriaceae;		
	Mycobacterium; Mycobacterium tuberculosis complex.		
REFERENCE			
AUTHORS	Imboden, P., Troller, R., Marchesi, F., Telenti, A., Bodmer, T., Cole, S., Schopfer, K. and Burkart, T.		
TITLE	The rpoB gene of Mycobacterium tuberculosis		
JOURNAL	Unpublished		
REFERENCE			
AUTHORS	Imboden, P.		
TITLE	Direct Submission		
JOURNAL	Submitted (11-JUL-1994) Paul Imboden, Institute for Medical Microbiology, University of Berne, Friedbuehlstrasse 51, Berne, 3010, Switzerland		
FEATURES			
source	Location/Qualifiers		
	1. .3853		
	/organism="Mycobacterium tuberculosis"		
	/strain="H37Rv"		
	/db_xref="taxon:1773"		
gene	576..>3853		
	/gene="rpoB"		
CDS	576..>3853		
	/gene="rpoB"		
	/codon_start=1		
	/transl_table=11		
	/product="RNA-polymerase beta subunit"		
	/protein_id="AA20242.2"		
	/db_xref="GI:714499"		
	/translation="MLEGCILADRSQKTAASPSRPSQSSNNNSVPGAPNRVSFAKL		
	REPLEVPLGLDVQDSFEWLGSPRMSAAERGVNVPVGLFEVLYELSPEDFSGS		
	MSLSFSDPRDDVKAPVDECKDKDMYAAPLFVTAEFINNNTGEIKSVTFVMDPDM		
	TEKGTITNTERVVSOLVRSPGVFDETDKSTDKTLHSVKVIPSRCAMLEFDVK		
	RDVGVIRKRRQPTVTLKALGVTSQIVRFGSEIHRSTLEKLVHGTVEALLD		
	IYKRLRGPDPPTKESQTLLENLFFKEKRYDLARVGRYKVNKKLGHVGEPTTSFLT		
	EEDVWATIEYLVRLHREGQTTMTVPGGVEVPVETDDIDHFGNRRRLRTVGLIQNIRVG		

MSRERVVRERWTTQDYEAITPOTLINIRPVVAIKEFFCTGTSOLSFQMDNNPLSGLT			
HKRLSALPGGLSRERAGLVDVDFHSHYGRMCPIETPGPNIGLIGLSVYARVNP			
FGFIIPYRKVVDGVSDGIVILTAEEDRRHVAQANSPIDADGRFVEPVLVRRKAP			
EVEYVPSVEDYMDVSPRWVATAMIPLEHDDANRMLGMORQAVPLVRSEAP			
LVGTGMELRAIDAIDATSSOSBSEVIEVSADYITVMDHNGTTRTYRNRKFAARNHGT			
ANOCPIVDAGDRVAGOVADGCTDDEMGALKNLLVAIMPHEGHNYDAIILSNRL			
VEEDVLSIHTEHEHIDARDPKLGAERTIDIPNISDEVLAIDLDERGIVRTGARVDC			
DILYGVTPKGETELTPEERLLRAIFGEKAREVRDTSIKVPHGSEGVIGIRVPSRED			
EDELPGAVNELRVYVAVKRWISDGLAGHNGKVGKILPVEDLDPFLADGTPVDI			
ILNTHGPRRRNIGIILETHLGWCAHSGWKVDAAGVYDWAARLPDELLAQPNATVS			
TPVFDGAQEAELQGLLCTLPNRDGDVLVDADGKAMLFDRSGSEFPYPVTVGYMYIM			
KLHLVDDKIHARSTGYPYSMITQOPLGKAQGGQRFGECHWANOAYGAAYTLQELL			
TIKS"			
BASE COUNT	723 a 1173 c 1293 g 664 t		
ORIGIN			
Query Match	80.0%;	Score 166.4;	DB 1; Length 3853;
Best Local Similarity	87.5%;	Pred. No. 1e-16;	
Matches 182;	Conservative 0;	Mismatches 26;	Indels 0; Gaps 0;
Qy 1	TCAAGGAGAAGCGCTAGCAGCTGGCCCGGTAGCCGCTACAGGTCACAAAGAAGCTCG	60	
Db 1477	TCAAGGAGAAGCGCTAGCAGCTGGCCCGGTAGCCGCTATTAAGGTCAACAGAAGCTCG	1536	
Qy 61	GCCTGCAGCTCGCGATCCGATCACCAGCTCCACGCTGACCGAGAGAGCTCGTCGCCCA	120	
Db 1537	GGCTGCATGTCCGGAGGCCATCAGCTCGCTGCGTATTAAGGTCAACAGAAGCTCG	1596	
Qy 121	CCATCGAGTACCTGCTCCGCTGACGAGGCCAGCACGATGACCGTCCCGGGCGGCA	180	
Db 1597	CCATCGAATATCTGCTCCGCTTGCAGGAGGTGACAGCACGATGACCGTCCCGGGCGGCG	1656	
Qy 181	CCGAGGTCCCGTTGAGACCGACGACAT	208	
Db 1657	TCGAGGTCCCGTGGAAACCGACGACAT	1684	
RESULT 12			
MSGRPOB			
LOCUS	5084 bp	DNA	linear BCT 13-SEP-1994
DEFINITION	Mycobacterium tuberculosis RNA polymerase beta-subunit (rpoB) gene, complete cds and RNA polymerase beta'-subunit rpoC gene, partial cds.		
ACCESSION	L27989		
VERSION	L27989.1	GI:468333	
KEYWORDS	RNA polymerase beta-subunit; rpoB gene.		
SOURCE	Mycobacterium tuberculosis (strain Rv)		
ORGANISM	Mycobacterium tuberculosis		
	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;		
	Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium		
	tuberculosis complex.		
REFERENCE			
AUTHORS	Miller, L.P., Crawford, J.T. and Shinnick, T.M.		
TITLE	The rpoB gene of Mycobacterium tuberculosis		
JOURNAL	Antimicrob. Agents Chemother. 38 (4), 805-811 (1994)		
MEDLINE	94304130		
PUBMED	8031050		
FEATURES			
source	Location/Qualifiers		
	1. .5084		
	/organism="Mycobacterium tuberculosis"		
	/strain="Rv"		
	/db_xref="taxon:1773"		
gene	1065..4598		
	/gene="rpoB"		
CDS	1065..4598		
	/gene="rpoB"		
	/codon_start=1		
	/evidence=experimental		
	/transl_table=11		
	/product="RNA polymerase beta-subunit"		
	/protein_id="AAA21416.1"		
	/db_xref="GI:468334"		
	/translation="MLEGCILADRSQKTAASPSRPSQSSNNNSVPGAPNRVSFAKL		





SPASGRPLAMPRLDMVTGLVYLTTTEVPEDTGEYOPASGDHPETGVYSSPAEAIMAADR  
GVLVRAKIVRLTQLRPPVEIEAELEFGSGMQPGDAMWAETTLGRVMEHELLPLGYP  
FVNQKHKKVQAAIINDLAERYPMIVVAQTVKDLKAGFYWATRSVTVSMADVLVPP  
RKEILDHYEERADKVEKQFQALNHERNEALVEIKWEATDEVEGQALREHYPPDNP  
IITIVSDGNGTQRTLAGMKGLVTPKGEFIPRPVKSSEFREGTLVLEYFINTHGA  
RKLADALTADSGYLTRLRLVDVSDQVIVREHDCOTERGIIVVELAERAPDGLIRDP  
YIETSAVARTGLTDAVEAGNIVVERGQDGDPEIDALLAAGITOVKRVSLTCAITST  
GVATCYGRSMATGKLVLDIGEAVGIVAAGSIGEPGTQMTMRTFHOGGVGDCITGGLPR  
VOELFEARPRGKAPLADTVGRVRLDGERFYKITIVPDDGEEVGYDXIKSRQRLRV  
FKHEDSERVLSGDHVEVGOQLMEGSADPHEVLRVQGPREVQIHLVRVQEVYRAQG  
VSLHDKHIEVILQMLRRVITIDSGSTFELPGLSLDRCAFEAENRRVVAEGGEPAAGR  
PVLHGHTKSLATSLVLSAASFQETTRVLDAAINRSDKLNKLENVITIGKLIPAGT  
GINRYNIAVQPTTEARAAYTIPSYEDQYISPDFGAATCAAVPLDDYGSYDR"  
/gene  
CDS  
/complement(7691..8065)  
/gene="MT0697"  
/complement(7691..8065)  
/gene="MT0697"  
/note="identified by Glimmer2; putative"  
/codon\_start=1  
/transl\_table=1  
/product="hypothetical protein"  
/protein\_id="AAK44923.1"  
/db\_xref="GI:13880220"  
/translation="MEDSAAAI7NPGHAMASAMERSGLLECVAAGLDEQPFGEFTADKL  
NPDGSSRRVPRRQADGATGATHVERGGQSGQAGVVPQRMHGFPALAMQDRLIHH  
GEQTQNRIAQAFVRFCVCSPT"  
complement(8058..9572)  
/gene="MT0698"  
/note="This region contains an authentic frame shift and  
is not the result of a sequencing artifact; identified by  
Glimmer2; putative; conserved hypothetical protein,  
authentic frameshift"  
10167..10925  
/gene="MT0699"  
10167..10925  
/gene="MT0699"  
/note="identified by match to PFAM protein family HMM  
PF01261"  
/codon\_start=1  
/transl\_table=1  
/product="AP endonuclease, family 2"  
/protein\_id="AAK44924.1"  
/db\_xref="GI:13880221"  
/translation="MLIGSHVSTPDLAAAEAGADVQVIFLGNPQSKAPKPRDDAA  
ALKAATLPIVYHAPYLINLASANNRVRIPSRKILQETCAAAADIGAAAVIVHGHVAD  
DNIDGFGFWRKALDRLETPVYVLENTAGDGHANARREDTIALRWLDVIGDTGIFGC  
LDYCHTWAGEALTDAVRIKAITGRIDLHVCHDNRDEAGSGRDRHANLGSQIDPDL  
LVAARVAAAPVICETADQGRKDDIAFLRITGS"  
10957..11799  
/gene="MT0700"  
10957..11799  
/gene="MT0700"  
/note="similar to GB:U00012 PID:466863; identified by  
sequence similarity; putative"  
/codon\_start=1  
/transl\_table=1  
/product="hydrolase/esterase, putative"  
/protein\_id="AAK44925.1"  
/db\_xref="GI:13880222"  
/translation="MLRRVAILLAVALAFAGSGGTRRLAAGFGNGNSVHTLVDGAGR  
SYRLKPVGLPSSAPLVMHLGGFGSAKQERSYGWDELASEKELVAYPDGYHRAWN  
ANGGCGGREGVDDIGFVRAVVADIANNYSIDPARVYVTGMSNGAIMSYTLACNT  
SIFAAIIVSGTQDPCQSPRPVSVIHIGTADPLVRYHGGPCAGFARIDPPVPDLN  
AFWRVNRGALDTTTEGPTTSGATCADNRRVLLTVDDAGHRWPSFATQTLWRFFA  
AHRF"  
11859..13487  
/gene="MT0701"  
11859..13487  
/gene="MT0701"  
/note="similar to SP:P33224 GB:L20915 PID:457172  
PID:457174 PID:537028; identified by sequence similarity;  
putative"  
/codon\_start=1

/transl\_table=11  
/product="acyl-CoA dehydrogenase, putative"  
/protein\_id="AAK44926.1"  
/db\_xref="GI:13880223"  
/translation="MSDTHVTVNQVPPLENTNPASSPVLIEALIQEQGMGLDEVNEV  
GAISCAQWGLADNRNPILTHDAYGYRDEVEYDPAHELMRTAITHGHAAAP  
WADDPGAAHVRAAKTSVMTVEPGHICPISMTYAVVPALRYNSLAAYVPLLTSTREY  
DPELPAATKAGITAGNSMTKEQSGSVRAGTTOATPNADGSYSLTGHKKWTSAPMCD  
IFLVLAQAPDSCGLLPRVLDGTRNRNMFQRLKDLGNHNANASSEVEYDGAVALV  
GEEGVPTIITEMVNLRLDCLGASMTGLTRAVHHAHQHAKFAGAYLIDQPLMRN  
VLADIAVEAATIVAMRAGATDNVAGNETEALLRRIGLAAAKYVWCKRSTAAAE  
ALECLGGVYVEDSCMPRLYREAPLMGWEGSNVSALDTRAMATRPACVEVLFDEL  
ARSAGDPRLDGHVRLRPQLGDLDTIGYRARKTAEDICLALOGSLLVLRHGHGPAVAEA  
FLATRLGGWGAGYCTMPAGLDLAPILERALVKG"  
13498..14436  
/gene="MT0702"  
13498..14436  
/gene="MT0702"  
/note="similar to GP:3885480; identified by sequence  
similarity; putative"  
/codon\_start=1  
/transl\_table=11  
/product="enoyl-CoA hydratase/isomerase family protein"  
/protein\_id="AAK44927.1"  
/db\_xref="GI:13880224"  
/translation="MTHAIRVDFDNLKMTYEVTVGTRIARTFNRPKGNAILIADTPL  
ELSAVERADLDPGVHVILVSGREGFCAGFDLSAYAESSSTGGGAYQCTVLDGKT  
QAVNHLPNQMDPIDYQOMSRFVRGFSALMHADKPTVYKIHGYCVAGGTDIALHADO  
VIAADAKIGYPTTRVGVGPAAGLWHLRQDRAKRLIFTGDCITGAQAAEWGLAVEA  
PEPADLDERLVARIAALPNOLIMVKLALNSALLQOQVATSRMSTVFDGAARHT  
PEGHAFVADAVEHGFDAVRRDEPFQDYGQASRV"  
14439..15161  
/gene="MT0703"  
14439..15161  
/gene="MT0703"  
/note="identified by Glimmer2; putative"  
/codon\_start=1  
/transl\_table=11  
/product="hypothetical protein"  
Query Match 80.0%; Score 166.4; DB 1; Length 19352;  
Best Local Similarity 87.5%; Pred. No. 7.7e-17;  
Matches 182; Conservative 0; Mismatches 26; Indels 0; Gaps 0;  
Qy 1 TCAAGGAGAGCGCTACGACCTGGCCCGGTAGCGGTACCAAGGTCAACAAGAGCTCG 60  
|||||  
Db 1064 TCAAGGAGAGCGCTACGACCTGGCCCGGTAGCGGTATTAAGGTCAACAAGAGCTCG 1123  
|||||  
Qy 61 GCCTGCACGTCGCGCATCCGATCCAGCTCCACGTCACCGAGGAGAGCTCGTCGCCA 120  
|||||  
Db 1124 GCCTGCATGTCGGGAGCGCCATCAGCTGTCGACGTCAGCGAAGAGCTCGTGCCCA 1183  
|||||  
Qy 121 CCATCGAGTACCTGTCGCCCTGACGAGGCGCCAGCACGATGACCTGCCGCGCGGCA 180  
|||||  
Db 1184 CCATCGAATATCTGTCGCCCTTGCACGAGGCTCAGACCATGACCTGTCGCGCGCGC 1243  
|||||  
Qy 181 CCGAGTGCCCGTTGAGACCGGACGACAT 208  
|||||  
Db 1244 TCGAGTGCCCGTGTGAAACCGGACGACAT 1271  
|||||  
RESULT 14  
MTC1376  
LOCUS  
DEFINITION Mycobacterium tuberculosis H37Rv complete genome; segment 32/162.  
ACCESSION 295972 AL123456  
VERSION 295972.1 GI:3261790  
KEYWORDS  
SOURCE Mycobacterium tuberculosis H37Rv.  
ORGANISM Mycobacterium tuberculosis H37Rv  
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
Actinomycetales; Corynebacterineae; Mycobacteriaceae;  
Mycobacterium; Mycobacterium tuberculosis complex.



probablemembrane protein, similar to YPRB\_ECOLI\_P33774  
hypothetical24.3 kd protein (unf 1) (217 aa), fasta  
scores: opt: 174 z-score: 215.8 E(): 5.3e-05, (25.6%  
identity in 223 aa overlap). Also similar to MTCY359.10  
(28.7% identity in 178 aa overlap)"  
/codon\_start=1  
/transl\_table=11  
/product="hypothetical protein Rv0658c"  
/protein\_id="CAB09376.1"  
/db\_xref="GI:2143303"  
/db\_xref="SPTREMBL:O06781"  
/translation="MEAGRADTVAPSHRWGLGAFVLVFLVASTSLAVVLTHGVP  
SAGVLALAAPTVVAGLAIIITRLRNGRLTDLRLRWSRGLRLGLMFGFGMLVT  
IPASLVYTAIVGPENSAVRIFGVRASWPALVFLVVFVAPLCEIIIRGLLNG  
AVDRWGMALVTVTVFALAHLEFARAPLLVVAIPALAREFYSGLLASIVTHQV  
TNLPGIVLLGLTGAISLP"  
complement(4480..4483)  
/note="possible RBS upstream of Rv0658c"  
complement(4753..5061)  
/gene="Rv0659c"  
complement(4753..5061)  
/gene="Rv0659c"  
/note="Rv0659c", (MTCI376.17), len: 102; unknown, similar  
to YW28\_MYCTU\_Q10867 hypothetical 12.3 kd protein cy39.28  
(114 aa), fasta scores: opt: 144 z-score: 213.2 E():  
7.3e-05, 30.8% identity in 107 aa overlap. Also similar to  
MTCY09f9.22 (32.7% identity in 101 aa overlap)"  
/codon\_start=1  
/transl\_table=11  
/product="hypothetical protein Rv0659c"  
/protein\_id="CAB09387.1"  
/db\_xref="GI:2143302"  
/db\_xref="SPTREMBL:O06780"  
/translation="MRGELWFAATPGGDRPVLVLTDPVADRGVAVVVALTRTRRG  
LVSELELAVENRVPSDCVNFNIHTLPRTAFRRIRITLSPARLHEACQTLRASTGC  
complement(5048..5293)  
/gene="Rv0660c"  
complement(5048..5293)  
/gene="Rv0660c"  
/note="Rv0660c", (MTCI376.16), len: 81, some similarity to  
IAF016485\_130 Halobacterium sp: NRC-1 plasm (100 aa),  
32.4%identity in 74 aa overlap"  
/codon\_start=1

Query Match 80.0%; Score 166.4; DB 1; Length 19770;  
Best Local Similarity 87.5%; Pred. No. 7.7e-17;  
Matches 182; Conservative 0; Mismatches 26; Indels 0; Gaps 0;  
Qy 1 TCAAGGAGAGCGCTACGACCTGGCCGGGTAGCGGCTACAGGCTCAACAGAGAGCTCG 60  
Db 10758 TCAAGGAGAGCGCTACGACCTGGCCGGGTAGCGGCTATTAAGGTCAACAGAGAGCTCG 10817  
Qy 61 GCCTGCACGTCGCGGATCGGATCCAGCTCCACGCTGACCCGAGGAGAGCTCGCGCA 120  
Db 10818 GGCTGCATGTCGCGAGGCCATCAGCTCGTCGACGCTGACCCGAGAGAGCTCGTGCGCA 10877  
Qy 121 CCATCAGTACCTGTCGCGCTGCAGAGGCCACGACACAGATGACCGTCCCGGGCGGCA 180  
Db 10878 CCATCGAATATCTGTCGCTTGCAGAGGGTCAGACACGATGACCGTTCGCGGGCGG 10937  
Qy 181 CCGAGTCCCGGTTGAGACCGAGCAT 208  
Db 10938 TCAGAGTCCCGTGGAAACCCGAGCAT 10965

RESULT 15  
MSU24494  
LOCUS MSU24494 3752 bp DNA linear BCT 02-MAR-2000  
DEFINITION Mycobacterium smegmatis DNA polymerase (rpoB) gene, complete cds.  
ACCESSION U24494  
VERSION U24494.1 GI:790347  
KEYWORDS

SOURCE  
ORGANISM Mycobacterium smegmatis.  
Mycobacterium smegmatis  
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
Actinomycetales; Corynebacterineae; Mycobacteriaceae;  
Mycobacterium.  
REFERENCE 1 (bases 1 to 3752)  
Hetherington,S.V., Watson,A.S. and Patrick,C.C.  
Sequence and analysis of the rpoB gene of Mycobacterium smegmatis  
Antimicrob. Agents Chemother. 39 (9), 2164-2166 (1995)  
MEDLINE 96050766  
PUBMED 8540740  
REFERENCE 2 (bases 1 to 3752)  
Hetherington,S.V.  
Direct Submission  
Submitted (11-APR-1995) Seth V. Hetherington, Infectious Diseases,  
St. Jude Children's Research Hospital, 332 N. Lauderdale, Memphis,  
TN 38101, USA  
FEATURES  
Location/Qualifiers  
1..3752  
/organism="Mycobacterium smegmatis"  
/db\_xref="taxon:1772"  
194..3703  
/gene="rpoB"  
194..3703  
/gene="rpoB"  
/codon\_start=1  
/evidence=experimental  
/transl\_table=11  
/product="DNA polymerase"  
/protein\_id="AA91426.2"  
/db\_xref="GI:7144498"  
/translation="MLEGCILAVSSQSKSNATINNSVRGPNGVSFAKUREPLEVRGL  
LDVQTSFNLVGRSRRVQAADIRGEENPVGLLEAEISPIEDFSGSMSLSFSDPR  
FDVQTSDECKDMYAAPLFVTAETINNTGEIKSOTVFMGDFPMTEKGTFTIN  
GTERVVSOLVRSPGVYFDETDKSTKTLHSVKVTPGRGAMLEDDVDKRDVGVRI  
RKRPQVTLIKALGNTNQIIVERGFCSEIMGTLEKDTSGTDEALLDIYKLRPGE  
PPTKSAQTLLENLFKERRKIDLRVGRYKKNKGLNAGKPTTSTLTEDVYATIE  
YLRLHEGTSMTVPGVEFAVESHDIDHGNRLRTVGELEONQIRVLSRHRVVR  
ERMTTQDVEAITPTLINIRVVAAIKKEFGTSQSFQNDQNNPLSLGTHKRLSALG  
PGLSRERAGLEVVRVHSHYRMCPIETPEGNIGLISLSVYARVNPFGFTETPYR  
KVVEGVVTDIDLTADERDHRVAAQNSPTDENGFTEDRVVVRKKGGEVEFVSADQ  
VDYMDVSPQMYSVATMIPFLEHDDANRALCANQROAVPLVSEAPLVCTGHFLR  
AATDAVWSHKTGCTVIEVSADYITVHADDGTROSLRKLKFKANSNGTGANRPIVDA  
GQREAGQVIADGCTQNGEMALGNLLVAIMPHEGNYEDAILNSRLVEEDVLTSI  
HIEEHIDARDTKLGAEEITRVIPKLSDEVLADLPRIVRIGAEVYRDEILVQKVT  
KGTELTPEERLLAIFGEKAREVRDTSKVPHGSEKVGIRVFSREDDDELPGVGN  
ELVRYVVAQRKISDGDKLARRHNGKVGKILPVEDMPFLPDGTPVDIILNTHGCRV  
VLNIGOILETHLGMGAKMNIDVLAGVPDMASKLPEELYSAPADSTVATPVQGAQE  
GELAGLIGSTLPNRDGEVYVNDGKATLFDGRSSGEPFPYPTVGVYMYLKLHLVDD  
KHARSTGPTSMITQPLGKAQFGGREGMECWAMQAGYATILQELLTINSDDTV  
GRVYVEAIVKGENIPEGIPESFKYLLKELQSLNVLSSDGRAIEMRDGDDDL  
ERAAANLGINLSRESASVEDLA"  
BASE COUNT 706 a 1225 c 1210 g 611 t  
ORIGIN  
Query Match 76.9%; Score 160; DB 1; Length 3752;  
Best Local Similarity 85.6%; Pred. No. 1e-15;  
Matches 178; Conservative 0; Mismatches 30; Indels 0; Gaps 0;  
Qy 1 TCAAGGAGAGCGCTACGACCTGGCCGGGTAGCGGCTACAGGCTCAACAGAGAGCTCG 60  
Db 1068 TCAAGGAGAGCGCTACGACCTGGCCGGGTAGCGGCTTACAAGGTCAACAGAGAGCTCG 1127  
Qy 61 GCCTGCACGTCGCGGATCCGATCACCAGCTCCACGCTGACCGAGGAGAGAGCTCGTGCGCA 120  
Db 1128 GCCTGAACGCGGCAACCGCATCACCAGCTCGACGCTGACCGAGAGAGAGCTCGTGCGCA 1187  
Qy 121 CCATCAGTACCTGTCGCGCTGCACAGGGCCAGCAGACATGACCGTCCCGGGCGGCA 180  
Db 1188 CCATCAGTACCTGTCGCTGCACAGGGCTCAGACCTCGACCTCCCGGGTGGCG 1247  
Qy 181 CCGAGGTGCCGGTTGAGACCGAGCAT 208

Db 1248 TCGAGTTCGCGTCGAGTCGACGACAT 1275

Search completed: November 13, 2002, 01:25:04  
Job time : 675.723 secs

GenCore version 5.1.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: November 12, 2002, 15:56:36 : Search time 27.8936 Seconds  
(without alignments)  
2286.857 Million cell updates/sec

Title: US-09-697-123b-8  
Perfect score: 208  
Sequence: 1 tcaagagaagcgctacgac.....ccggtggaaccgacgacat 208

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents\_NA:\*

- 1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq:\*
- 2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq:\*
- 3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq:\*
- 4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq:\*
- 5: /cgn2\_6/ptodata/1/ina/PTUS\_COMB.seq:\*
- 6: /cgn2\_6/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	166.4	80.0	970	1	US-08-250-030-1
2	166.4	80.0	970	5	PCT-US95-06790-1
3	142.4	68.5	3447	2	US-08-313-185-57
4	142.4	68.5	3447	3	US-09-082-614A-57
5	39.4	18.9	1894	4	US-09-329-350-32
6	39.4	18.8	1100	3	US-09-248-335-53
7	38.8	18.3	2742	4	US-09-232-468A-1
8	38.8	18.3	4897	6	5196516-7
9	36.8	17.7	4411529	4	US-09-103-840A-1
10	36.2	17.4	2634	1	US-08-196-218-31
11	36.2	17.4	2634	1	US-08-681-953-31
12	35.8	17.2	734	4	US-09-221-017B-1070
13	34.8	16.7	4403765	4	US-09-103-840A-2
14	34.6	16.6	1366	4	US-09-280-116-120
15	34.2	16.4	20235	1	US-07-642-734C-3
16	34.2	16.4	20235	3	US-08-439-009A-3
17	34	16.3	1208	2	US-08-403-852D-4
18	34	16.3	1208	3	US-08-510-646B-4
19	34	16.3	1208	4	US-09-231-818-4
20	34	16.3	5392	2	US-08-403-852D-1
21	34	16.3	5392	3	US-08-510-646B-1
22	34	16.3	5392	4	US-09-231-818-1
23	33.8	16.2	6095	4	US-09-029-603-4
24	33.6	16.2	1272	3	US-09-036-987A-28
25	33.6	16.2	1272	4	US-09-370-700-28
26	33.4	16.1	1112	4	US-09-434-288-9
27	33.4	16.1	1281	4	US-09-105-537-19

c 28	33.4	16.1	13613	4	US-09-105-537-3	Sequence 3, Appl
c 29	33.4	16.1	30001	1	US-08-125-468-1	Sequence 1, Appl
c 30	33.4	16.1	30001	2	US-08-474-933-1	Sequence 1, Appl
31	33.4	16.1	38506	3	US-09-320-878-19	Sequence 19, Appl
32	33.4	16.1	44377	2	US-08-804-227C-7	Sequence 7, Appl
33	33.4	16.1	44377	2	US-08-804-198-1	Sequence 1, Appl
c 34	33.2	16.0	49272	1	US-08-614-770A-1	Sequence 1, Appl
35	33	15.9	765	3	US-08-718-904-79	Sequence 79, Appl
36	33	15.9	1722	4	US-09-385-028-15	Sequence 15, Appl
37	33	15.9	11604	4	US-09-385-028-13	Sequence 13, Appl
38	33	15.9	15079	4	US-09-385-028-1	Sequence 1, Appl
39	32.6	15.7	3048	1	US-08-188-228-47	Sequence 47, Appl
40	32.6	15.7	3048	1	US-08-332-643-41	Sequence 41, Appl
41	32.6	15.7	3048	1	US-08-332-638-47	Sequence 47, Appl
42	32.4	15.6	3435	1	US-08-366-577-1	Sequence 1, Appl
43	32.4	15.6	3435	5	PCT-US96-00005-1	Sequence 1, Appl
44	32.4	15.6	11219	1	US-07-642-734C-1	Sequence 1, Appl
45	32.4	15.6	11219	3	US-08-439-009A-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1  
US-08-250-030-1  
: Sequence 1, Application US/08250030  
: Patent No. 5643723  
: GENERAL INFORMATION:  
: APPLICANT: Persing, David H.  
: TITLE OF INVENTION: Detection of a Genetic Locus Encoding  
: TITLE OF INVENTION: Resistance to Rifampin in Mycobacterial Cultures and in  
: TITLE OF INVENTION: Clinical Specimens  
: NUMBER OF SEQUENCES: 15  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: Schwegman, Lundberg & Woessner  
: STREET: 3500 IDS Center  
: CITY: Minneapolis  
: STATE: MN  
: COUNTRY: USA  
: ZIP: 55402  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: COMPUTER: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: PatentIn Release #1.0, Version #1.25  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/08/250,030  
: FILING DATE: 26-MAY-1994  
: CLASSIFICATION: 435  
: ATTORNEY/AGENT INFORMATION:  
: NAME: Muetting, Ann M.  
: REGISTRATION NUMBER: 33,977  
: REFERENCE/DOCKET NUMBER: 150.105US1  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: 612-339-0331  
: TELEFAX: 612-339-3061  
: INFORMATION FOR SEQ ID NO: 1:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 970 base pairs  
: TYPE: nucleic acid  
: STRANDEDNESS: single  
: TOPOLOGY: linear  
: MOLECULE TYPE: DNA  
US-08-250-030-1

Query Match 80.0%; Score 166.4; DB 1; Length 970;  
Best Local Similarity 87.5%; Pred. No. 3.7e-33;  
Matches 182; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

OY 1 TCAAGGAGAGCGCTACGACCTGGCCCGGTGTGGCGGATACAGGTCAACAAGAGCTGG 60  
|||||  
DB 26 TCAAGGAGAGCGCTACGACCTGGCCCGGTGTGGCGGATACAGGTCAACAAGAGCTGG 85  
|||||



Sequence 57, Application US/09082614A  
Patent No. 6124098  
GENERAL INFORMATION:  
APPLICANT: Heym, Beate  
APPLICANT: Cole, Stewart  
APPLICANT: Young, Douglas  
APPLICANT: Zhang, Ying  
APPLICANT: Honore, Nadine  
APPLICANT: Telenti, Amalio  
APPLICANT: Rodmer, Thomas  
TITLE OF INVENTION: Rapid Detection of Antibiotic Resistance  
NUMBER OF SEQUENCES: 66  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &  
ADDRESSEE: Dunner  
STREET: 1300 I Street, N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20005-3315  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/082,614A  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/313,185  
FILING DATE: 12-OCT-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Meyers, Kenneth J.  
REGISTRATION NUMBER: 25,146  
REFERENCE/DOCKET NUMBER: 02356.0068-00000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 408-4000  
TELEFAX: (202) 408-4400  
INFORMATION FOR SEQ ID NO: 57:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3447 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-09-082-614A-57

Query Match 68.5%; Score 142.4; DB 3; Length 3447;  
Best Local Similarity 80.3%; Pred. No. 4.1e-27;  
Matches 167; Conservative 0; Mismatches 41; Indels 0; Gaps 0;  
QY 1 TCAAGGAGACGGCTACGACCTGGCCGCTCGGCCGATACAAAGTCAACGAAGCTGG 60  
Db 809 TCAAGGAGAACGGCTACGACCTGGCCGAGGTGGCTGTACAAAGGTCAACGAAGCTCG 868  
QY 61 GCCTGAACACCAATCATCCGATCACCACGACGCTGACCGAAGAGACGTCGTCGCCA 120  
Db 869 GGTTCACCGCGGTGAGTGTATCAGCTGCTCCAGCTGACCGAAGAGGATGTCGTCGCCA 928  
QY 121 CCATCGAGTATCTGGTCCGCTGACACGAGGCCAGGCCACGATGACCGTCCGGGGGGG 180  
Db 929 CCATAGAGTACCTGGTTCGTCCTCATGAGGTGAGTCCGACAACTGCTGCCAGGTGGG 988  
QY 181 TCGAGGTGCCGGTGGAAACCGACCAT 208  
Db 989 TAGAAGTCCAGTGGAAACTGACGATAT 1016

RESULT 5  
US-09-329-350-32  
Sequence 32, Application US/09329350

Patent No. 6184019  
GENERAL INFORMATION:  
APPLICANT: Miettinen-Oinonen, Arja  
APPLICANT: Lonsborough, John  
APPLICANT: Vehmaanper, Jari  
APPLICANT: Haakana, Hei  
APPLICANT: M ntyl , Arja  
APPLICANT: Lantto, Raija  
APPLICANT: Elovainio, Minna  
APPLICANT: Joutsjoki, Vesa  
APPLICANT: Paloheimo, Marja  
APPLICANT: Suominen, Pirkko  
TITLE OF INVENTION: NOVEL CELLULASES, THE GENES ENCODING THEM AND  
NUMBER OF SEQUENCES: 45  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.  
STREET: 1100 New York Avenue, N.W., Suite 600  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/329,350  
FILING DATE: Herewith  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/841,636  
FILING DATE: 30-APR-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/005,335  
FILING DATE: 17-OCT-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/007,926  
FILING DATE: 04-DEC-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/020,840  
FILING DATE: 28-JUN-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/732,181  
FILING DATE: 16-OCT-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/PI96/00550  
FILING DATE: 17-OCT-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Shea Jr., Timothy  
REGISTRATION NUMBER: 41,306  
REFERENCE/DOCKET NUMBER: 1716.0510006/MAC/TJS  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)371-2600  
TELEFAX: (202)371-2540  
INFORMATION FOR SEQ ID NO: 32:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1894 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
ORIGINAL SOURCE:  
ORGANISM: Melanocarpus albomyces  
STRAIN: ALK04237  
FEATURE:  
NAME/KEY: exon  
LOCATION: 233..838  
OTHER INFORMATION: /product= "50K-cellulase"  
FEATURE:  
NAME/KEY: exon  
LOCATION: 916..1596



```
; OTHER INFORMATION: /product- "50K-cellulase"
US-09-329-350-32

Query Match      18.9%; Score 39.4; DB 4; Length 1894;
Best Local Similarity 55.5%; Pred. No. 0.15;
Matches 76; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

U; 56 GCTGGGCTGAACACCAATATCCGATACACGACGACGCTGACCGAAGAAGACGCTCGT 115
   || || || || || || || || || || || || || || || || || || || || ||
Db 1131 GTTCCGCTGCACACGCGGCGCTTACCGTGTGTGACGCACTTCCCGCGCAGCCGGA 1190

QY 116 CGCCACCATCGAGTATCTGCTCGCCCTGACAGAGGCCAGGCCACCATGACCGTGCCGG 175
   || || || || || || || || || || || || || || || || || || || || ||
Db 1191 GGGCCGGCTCGAGAGCATCACCGCTGTACGTGCGAGGACGGCAAGGTATCGAGTCGTA 1250

QY 176 CGGGGTGAGGTGCCGG 192
   || || || || || || || || || || || || || || || || || || || || ||
Db 1251 CGTCGTGAGCGCCCGG 1267

RESULT 6
US-09-248-335-53
; Sequence 53, Application US/09248335
; Patent No. 6096504
; GENERAL INFORMATION:
; APPLICANT: MCGONIGLE, BRIAN
; APPLICANT: O'KEEF, DANIEL
; TITLE OF INVENTION: PLANT GLUTAMINOLINE-S-TRANSFERASE ENZYMES
; FILE REFERENCE: CL-1128-A
; CURRENT APPLICATION NUMBER: US/09/248,335
; EARLIER APPLICATION DATE: 1999-02-10
; EARLIER APPLICATION NUMBER: 08/924,759
; EARLIER FILING DATE: 1997-September-05
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: Microsoft Word Version 7.0A
; SEQ ID NO 53
; LENGTH: 1100
; TYPE: DNA
; ORGANISM: maize
US-09-248-335-53

Query Match      18.8%; Score 39; DB 3; Length 1100;
Best Local Similarity 52.8%; Pred. No. 0.18;
Matches 84; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 47 CAACAAGAAGCTGGGCTGAACACCAATATCTGCTCCGCTGACACGACGACGCTGACCGAAGA 106
   || | || || || || || || || || || || || || || || || || || || ||
Db 193 CAAGCGCGCTCTCTCTGGATCCAAACCCGCTGCAAGAGCGTCCCGTGTCTCTCTCCA 252

QY 107 AGACGCTGCTGCCACCATCCAGTATCTGCTCCGCTGACAGGCGCCAGCCACGATGAC 166
   || || || || || || || || || || || || || || || || || || || || ||
Db 253 CGCGCGCGCGCCATAACAGAGTCCCAGGTATCTCTGACATATCGACGAGGTGTGGGC 312

QY 167 CGTGCCGGCGGGCTGCGAGCTGCGGTGGGAACCGAGCA 205
   || || || || || || || || || || || || || || || || || || || || ||
Db 313 GGGGACGGCGCGCGCTGCTTCCGGCGGACCCCTAGCA 351

RESULT 7
US-09-232-468A-1
; Sequence 1, Application US/09232468A
; Patent No. 6207165
; GENERAL INFORMATION:
; APPLICANT: AUDONNET et al.
; TITLE OF INVENTION: POLYNUCLEOTIDE VACCINE FORMULA AGAINST PORCINE
; TITLE OF INVENTION: REPRODUCTIVE AND RESPIRATORY PATHOLOGIES
; FILE REFERENCE: 454313-2230
; CURRENT APPLICATION NUMBER: US/09/232,468A
; CURRENT FILING DATE: 1999-01-05
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2742
```

```
; TYPE: DNA
; ORGANISM: Pseudorabies virus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2742)
US-09-232-468A-1

Query Match      18.3%; Score 38; DB 4; Length 2742;
Best Local Similarity 60.8%; Pred. No. 0.36;
Matches 62; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 62 CTTGAACACCAATATCCGATCACGACGACGCTGACCGAAGAAGACGCTCGTCGCCAC 121
   || || || || || || || || || || || || || || || || || || || || ||
Db 2484 CATGAAGGCCCTGTACCCCGTCCAGGACGCTCAAGGAGGACGCGCTCGACGAAGG 2543

QY 122 CATCGAGTATCTGGTCCGCTGCAGGAGGCCAGGCCACGAT 163
   || || || || || || || || || || || || || || || || || || || || ||
Db 2544 CGACGTGGAGGAGGCCAAGCTGGACACGCCCGGGACATGAT 2585

RESULT 8
5196516-7
; Patent No. 5196516
; APPLICANT: SCHREURS, CHRISTA S.;METTENLEITER, THOMAS C.
; SIMON, ARTUR J.;LUKAS, NOEMI;RZIHA, HANNS J.
; TITLE OF INVENTION: PSEUDORABIES VIRUS VACCINE
; NUMBER OF SEQUENCES: 8
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/383,833
; FILING DATE: 21-JUL-1989
; SEQ ID NO:7;
; LENGTH: 4897
5196516-7

Query Match      18.3%; Score 38; DB 6; Length 4897;
Best Local Similarity 60.8%; Pred. No. 0.39;
Matches 62; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 62 CTTGAACACCAATATCCGATCACGACGACGCTGACCGAAGAAGACGCTCGTCGCCAC 121
   || || || || || || || || || || || || || || || || || || || || ||
Db 3355 CATGAAGGCCCTGTACCCCGTCCAGGACGCTCAAGGAGGACGCGCTCGACGAAGG 3414

QY 122 CATCGAGTATCTGGTCCGCTGCAGGAGGCCAGGCCACGAT 163
   || || || || || || || || || || || || || || || || || || || || ||
Db 3415 CGACGTGGAGGAGGCCAAGCTGGACACGCCCGGGACATGAT 3456

RESULT 9
US-09-103-840A-1/c
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

Query Match      17.7%; Score 36.8; DB 4; Length 4411529;
Best Local Similarity 55.5%; Pred. No. 1.9;
Matches 71; Conservative 0; Mismatches 57; Indels 0; Gaps 0;
```

```
QY 50 CAGAGAGCTGGCCCTGACACCAATCATCCGATCACCACGACGAGCTGACCGGAAGA 109
|| || || || || || || || || || || || || || || || || || || || || ||
Db 595927 CACCAGATGACCTGGACCGCGGCCAGCAGGTATGGCTGATCACCCTACCCCATACGA 595868
|| || || || || || || || || || || || || || || || || || || || || ||
QY 110 GCTGCTGCCACCACTGAGTATCTGGTCCGCTGCACGAGGCGCCAGCCAGATGACCGT 169
|| || || || || || || || || || || || || || || || || || || || || ||
Db 595867 ACTGCGGCCACCATGCCGCCCGCGCTCGGCTAATCGGCGCCCTGGGACCGCTGCCCGA 595808
|| || || || || || || || || || || || || || || || || || || || || ||
QY 170 GCGGCGCG 177
|| || || || || || || || || || || || || || || || || || || || || ||
Db 595807 GTCTGCTG 595800
|| || || || || || || || || || || || || || || || || || || || || ||

RESULT 10
US-08-196-218-31
: Sequence 31, Application US/08196218
: Patent No. 5614619
: GENERAL INFORMATION:
: APPLICANT: Piepersberg, Wolfgang
: APPLICANT: Stockmann, Michael
: APPLICANT: Taleghani, Kampiz Mansouri
: APPLICANT: Distler, Jurgen
: APPLICANT: Grabley, Susanne
: APPLICANT: Sichel, Petra
: APPLICANT: Brau, Barbara
: TITLE OF INVENTION: Secondary-Metabolite Biosynthesis Genes
: TITLE OF INVENTION: From Actinomycetes, Method of Isolating Them, and Their
: NUMBER OF SEQUENCES: 34
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
: ADDRESSEE: Dunner
: STREET: 1300 I Street, N.W.
: CITY: Washington
: STATE: D.C.
: COUNTRY: United States
: ZIP: 20005-3315
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/196,218
: FILING DATE: 25-AUG-1994
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Ogden, Stasia L.
: REGISTRATION NUMBER: 36,228
: REFERENCE/DOCKET NUMBER: 02481.1372-00000
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 202-408-4000
: TELEFAX: 202-408-4400
: INFORMATION FOR SEQ ID NO: 31:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2634 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 3..401
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 416..1531
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 1561..2625
: US-08-196-218-31
```

Query Match 17.4% Score 36.2; DB 1; Length 2634;

```
Best Local Similarity 49.2%; Pred. No. 1;
Matches 95; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

QY 12 CGGTACGACCTGGCCCGCTGTCGGCCGATACAAGGTCAACAAGAAGCTGGCGCTGAACACC 71
||||| || || || || || || || || || || || || || || || || || || || || || ||
Db 2197 CTCTACGACCGCTACTGGAAGGACACCGGGGAGGTGCGAGGAGCTCTTGAGTGAACAGC 2256
||||| || || || || || || || || || || || || || || || || || || || || || ||
QY 72 AATCATCCGATCACCAGACGACGCTGACGGAAGAAGACGCTGTCGCCACCATCGAGTAT 131
|| || || || || || || || || || || || || || || || || || || || || ||
Db 2257 CAGCTCTGACGCGCTGACCGCGCGCTGACGAGGACAGGTGACGCGGACAGCGTCTC 2316
||||| || || || || || || || || || || || || || || || || || || || || || ||
QY 132 CTGGTCCGCTGCACGAGGCGCCAGGCTACCATGACCTGTCGGCGCGGCGGTGCGAGTCTCG 191
||||| || || || || || || || || || || || || || || || || || || || || || ||
Db 2317 GTCGGCGGGTGTGATCGAGCGCGGGCGCGCATCTGCGGGTTCGAGGGCGCG 2376
||||| || || || || || || || || || || || || || || || || || || || || || ||

RESULT 11
US-08-681-953-31
: Sequence 31, Application US/08681953
: Patent No. 5710032
: GENERAL INFORMATION:
: APPLICANT: Piepersberg, Wolfgang
: APPLICANT: Stockmann, Michael
: APPLICANT: Taleghani, Kampiz Mansouri
: APPLICANT: Distler, Jurgen
: APPLICANT: Grabley, Susanne
: APPLICANT: Sichel, Petra
: APPLICANT: Brau, Barbara
: TITLE OF INVENTION: Secondary-Metabolite Biosynthesis Genes
: TITLE OF INVENTION: From Actinomycetes, Method of Isolating Them, and Their
: NUMBER OF SEQUENCES: 34
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
: ADDRESSEE: Dunner
: STREET: 1300 I Street, N.W.
: CITY: Washington
: STATE: D.C.
: COUNTRY: United States
: ZIP: 20005-3315
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/681,953
: FILING DATE: 30-JUL-1996
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/196,218
: FILING DATE: 25-AUG-1994
: ATTORNEY/AGENT INFORMATION:
: NAME: Ogden, Stasia L.
: REGISTRATION NUMBER: 36,228
: REFERENCE/DOCKET NUMBER: 02481.1372-00000
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 202-408-4000
: TELEFAX: 202-408-4400
: INFORMATION FOR SEQ ID NO: 31:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2634 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 3..401
```

```
;
;
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 416...1531
;
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1561...2625
;
US-08-681-953-31

Query Match
Best Local Similarity 17.4%; Score 36.2; DB 1; Length 2634;
Matches 95; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

QY 12 CGCTAGCAGCTGCCCGTCTGCGCGCATACAGGTCAACNAGAAAGCTGGCCCTGAACACC 71
Db 2197 CTACGACCGCTACTGGAAGGACACCGGAGCGTCCAGGACGCTCTTGAGTGCAACAGC 2256

QY 72 AATCATCCGATCACCGACGACGCTGACCGAAGACAGCTCTCGCCACCATCTGAGTAT 131
Db 2257 CACCTCTGACGCGCTGACCGCGCGCTGCGGACGAGCTCGACGCGGACGCGTCTC 2316

QY 132 CTGCTCCGCTGACGAGGCGCCAGGACGATGACCGTCTCCGCGCGGTCTGAGGTCCG 191
Db 2317 CTGCGCGCGCTGATGATGAGGCGCGCGCGCATCTCTCGGCTCGAGGCGCGC 2376

QY 192 GTCGAACCAGC 204
Db 2377 GCGATCATCGCGC 2389

RESULT 12
US-09-221-017B-1070
; Sequence 1070, Application US/09221017B
; Patent No. 6444799
;
; GENERAL INFORMATION:
; APPLICANT: ROSS, Bruce C.
; TITLE OF INVENTION: P. GINGIVALIS NUCLEOTIDES AND USES THEREOF
; NUMBER OF SEQUENCES: 1120
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FORSTER
; STREET: 755 PAGE MILL ROAD
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/221,017B
; FILING DATE: 23-DEC-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PP1182
; FILING DATE: 31-DEC-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PP1546
; FILING DATE: 30-JAN-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PP2911
; FILING DATE: 09-APR-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/AU98/01023
; FILING DATE: 10-DEC-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Monroy, Gladys H
; REGISTRATION NUMBER: 32,430
; REFERENCE/DOCKET NUMBER: 27340-20021.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-813-5600
; TELEFAX: 650-494-0792
; TELEX: 706141

;
; INFORMATION FOR SEQ ID NO: 1070:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 734 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: UNKNOWN
; ORIGINAL SOURCE:
; ORGANISM: PORYPHYROMONAS GINGIVALIS
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1...734
;
US-09-221-017B-1070

Query Match
Best Local Similarity 17.2%; Score 35.8; DB 4; Length 734;
Matches 80; Conservative 0; Mismatches 47; Indels 6; Gaps 1;

QY 1 TCAAGGAGAAGCGCTACGACCTGGCCCGTGTGCGCGCATACAAAGGTCAACAAGAAGCTGG 60
Db 4 TCTCTGATAAAGCATACGACTTGGCGGATGTCGGACGCTATCTGATCAATAAGAACTCA 63

QY 61 GCCTGAACACCAATCATCCGATCACACGACGACGCTGACCGAAGAAGACGCTGTCGCGCA 120
Db 64 ATTTGAATATTGATCTCTGATATC-----AAAGTCTGACCAATGAAGATATTATTGACA 117

QY 121 CCATCGAGTATCT 133
Db 118 TCATCAAGTATCT 130

RESULT 13
US-09-103-840A-2/c
; Sequence 2, Application US/09103840A
; Patent No. 6294328
;
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
;
US-09-103-840A-2

Query Match
Best Local Similarity 16.7%; Score 34.8; DB 4; Length 4403765;
Matches 71; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

QY 50 CAAGAAGCTGGCGCTGAACACCAATCATCCGATCACACGACGACGCTGACCGAAGAAGA 109
Db 597288 CACCCAGATGCACCTGGACGCGCGCAGGATATGGTGTATCACCCTACCCCATACGA 597229

QY 110 CCTCGTCGCCACCATCGAGTATCTGTCGGCTGCACGAGGCGCCAGGCACGATGACCGT 169
Db 597228 ACTCGCGGCCACCATCGCGCCCGCTCGGCTAACCGCGCCCTGGGCACGCTGCGCGA 597169

QY 170 GCCGGGCG 177
```



```
; OTHER INFORMATION: /function= "approximate span
; FEATURE: beta-ketoreductase of module 4"
; NAME/KEY: misc_feature
; LOCATION: 10225..10483
; OTHER INFORMATION: /function= "approximate span of
; OTHER INFORMATION: acyl carrier domain of module 4"
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 10723..20235
; OTHER INFORMATION: /codon_start= 10723
; OTHER INFORMATION: /function= "gene-eryA"
; OTHER INFORMATION: /product= "orf3 encoding modules 5 & 6
; OTHER INFORMATION: 6-deoxyerythronolide B formatio"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 10723..15165
; OTHER INFORMATION: /function= "approximate span of
; OTHER INFORMATION: module 5"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 10831..12174
; OTHER INFORMATION: /function= "approximate span of
; OTHER INFORMATION: beta-ketoacylACPSynthase domain of modul"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 12379..13350
; OTHER INFORMATION: /function= "approximatr span of
; OTHER INFORMATION: acyltransferase domain of module 5"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 14062..14610
; OTHER INFORMATION: /function= "approximate span of
; OTHER INFORMATION: beta-ketoreductase of module 5"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 14857..15114
; OTHER INFORMATION: /function= "approximate span of
; OTHER INFORMATION: acyl carrier domain of module 5"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 15166..20235
; OTHER INFORMATION: /function= "approximate span of
; OTHER INFORMATION: module 6"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 15172..16569
; OTHER INFORMATION: /function= "approximate span of
; OTHER INFORMATION: beta-ketoacylACPSynthase domain of modul"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 16768..17721
; OTHER INFORMATION: /function= "approximate span of
; OTHER INFORMATION: acyltransferase domain of module 6"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 18379..18921
; OTHER INFORMATION: /function= "approximate span of
; OTHER INFORMATION: beta-ketoreductase domain of module 6"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 19149..19398
; OTHER INFORMATION: /function= "approximate span of
; OTHER INFORMATION: acyl carrier domain of module 6"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 19492..20235
; OTHER INFORMATION: /function= "approximate span of
; OTHER INFORMATION: thioesterase domain of module 6"
; US-07-642-734C-3
```

Query Match 16.4%; Score 34.2; DB 1; Length 20235;  
Best Local Similarity 52.4%; Pred. No. 4.2;

	Matches	75;	Conservative	0;	Mismatches	68;	Indels	0;	Gaps	0;
Oy	64	TGAACACCAATCATCCGATCACCACGACGAGCTGACCGAAGAAGACGCTCGTCGCCACCA	123							
Db	17573	TGAGCCCGCACCCGGTGTCTACCGCGGGGGTGCAGGAGATCGCCGGGACGCCGTGGCCA	17632							
Oy	124	TCGAGTATCTGGTCCGCTGCACGAGGCCAGGCCACGATGACCGTGCCGGGGGGTFCG	183							
Db	17633	TCGGTCTGCTGCACCGCGACACCGGAGGAGCACCTGATCGCGGAGCTCGCCGGGCGC	17692							
Oy	184	AGGTGCCCGGTGAAACCGACGAC	206							
Db	17693	ACGTGCACGGCTGCCCGTGGAC	17715							

Search completed: November 12, 2002, 20:14:08  
Job time : 5353.89 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 12, 2002, 16:05:40 ; Search time 21.1607 Seconds  
(without alignments)  
3487.380 Million cell updates/sec

Title: US-09-697-123b-8  
Perfect score: 208  
Sequence: 1 lcaaggagaagcgtacgac.....ccggtggaaccagcagcat 208

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 320260 seqs, 177392727 residues

Total number of hits satisfying chosen parameters: 640520

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications, NA:\*

- 1: /cgn2\_6/ptodata/2/pubpna/us07\_PUBCOMB.seq.\*
- 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*
- 3: /cgn2\_6/ptodata/2/pubpna/us06\_NEW\_PUB.seq.\*
- 4: /cgn2\_6/ptodata/2/pubpna/us06\_PUBCOMB.seq.\*
- 5: /cgn2\_6/ptodata/2/pubpna/us07\_NEW\_PUB.seq.\*
- 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq.\*
- 7: /cgn2\_6/ptodata/2/pubpna/us08\_NEW\_PUB.seq.\*
- 8: /cgn2\_6/ptodata/2/pubpna/us08\_PUBCOMB.seq.\*
- 9: /cgn2\_6/ptodata/2/pubpna/us09\_NEW\_PUB.seq.\*
- 10: /cgn2\_6/ptodata/2/pubpna/us09\_PUBCOMB.seq.\*
- 11: /cgn2\_6/ptodata/2/pubpna/us10\_NEW\_PUB.seq.\*
- 12: /cgn2\_6/ptodata/2/pubpna/us10\_PUBCOMB.seq.\*
- 13: /cgn2\_6/ptodata/2/pubpna/us60\_NEW\_PUB.seq.\*
- 14: /cgn2\_6/ptodata/2/pubpna/us60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	80	38.5	5096	10 US-09-984-711-5	Sequence 5, Appli
2	80	38.5	5099	9 US-10-075-460-5	Sequence 5, Appli
3	80	38.5	5099	10 US-09-887-052-1	Sequence 1, Appli
4	80	38.5	5099	10 US-09-887-052-3	Sequence 3, Appli
5	80	38.5	5099	10 US-09-887-052-5	Sequence 5, Appli
6	37.6	18.1	704	12 US-10-062-254-109	Sequence 109, App
7	37.6	18.1	752	12 US-10-062-254-111	Sequence 111, App
8	37.6	18.1	897	10 US-09-815-242-4001	Sequence 4001, Ap
9	37	17.8	492	10 US-09-815-242-7946	Sequence 7946, Ap
10	36.2	17.4	278	10 US-09-294-093B-1392	Sequence 1392, Ap
11	36.2	17.4	2693	10 US-09-880-107-3678	Sequence 3678, Ap
12	35.6	17.1	1437	10 US-09-815-242-7738	Sequence 7738, Ap
13	34.8	16.7	284	10 US-09-864-761-3324	Sequence 3324, A
14	34.8	16.7	569	10 US-09-864-761-15816	Sequence 15816, A
15	34.8	16.7	622	10 US-09-923-779-43	Sequence 43, Appli
16	34.8	16.7	1266	10 US-09-815-242-7920	Sequence 7920, Ap
17	34.8	16.7	1640	10 US-09-925-300-229	Sequence 229, App
18	34.8	16.7	3522	10 US-09-815-242-7837	Sequence 7837, Ap
19	34	16.3	4863	10 US-09-815-242-4071	Sequence 4071, Ap

20	33.6	16.2	939	10 US-09-815-242-4158	Sequence 4158, Ap
21	33.4	16.1	234	10 US-09-923-876-3102	Sequence 3102, Ap
c 22	33.4	16.1	709	9 US-09-992-598-280	Sequence 280, App
c 23	33.4	16.1	709	10 US-09-989-722-280	Sequence 280, App
c 24	33.4	16.1	709	10 US-09-989-722-280	Sequence 280, App
c 25	33.4	16.1	709	10 US-09-989-727-280	Sequence 280, App
c 26	33.4	16.1	709	10 US-09-989-727-280	Sequence 280, App
c 27	33.4	16.1	709	10 US-09-989-731-280	Sequence 280, App
c 28	33.4	16.1	709	10 US-09-989-732-280	Sequence 280, App
c 29	33.4	16.1	709	10 US-09-991-073-280	Sequence 280, App
c 30	33.4	16.1	709	10 US-09-990-442-280	Sequence 280, App
c 31	33.4	16.1	709	10 US-09-991-163-280	Sequence 280, App
c 32	33.4	16.1	709	10 US-09-993-604-280	Sequence 280, App
c 33	33.4	16.1	709	10 US-09-990-456-280	Sequence 280, App
c 34	33.4	16.1	709	10 US-09-989-721-280	Sequence 280, App
c 35	33.4	16.1	1281	10 US-09-861-289-19	Sequence 19, Appli
c 36	33.4	16.1	1665	10 US-09-815-242-7935	Sequence 7935, Ap
c 37	33.4	16.1	13613	10 US-09-861-289-3	Sequence 3, Appli
c 38	33.2	16.0	175	10 US-09-923-876-5702	Sequence 5702, Ap
c 39	33.2	16.0	2329	10 US-09-816-828-9	Sequence 9, Appli
c 40	33	15.9	862	10 US-09-733-569A-6	Sequence 6, Appli
c 41	33	15.9	1164	10 US-09-815-242-7696	Sequence 7696, Ap
c 42	32.8	15.8	43058	10 US-09-954-456-292	Sequence 292, App
c 43	32.8	15.8	43058	10 US-09-954-456-529	Sequence 529, App
c 44	32.8	15.8	43058	10 US-09-880-107-3950	Sequence 3950, Ap
c 45	32.6	15.7	626	12 US-10-062-254-115	Sequence 115, App

ALIGNMENTS

RESULT 1  
US-09-984-711-5  
; Sequence 5, Application US/09984711  
; Patent No. US20020119549A1  
; GENERAL INFORMATION:  
; APPLICANT: MOECKEL, Bettina  
; APPLICANT: BATHE, Brigitte  
; APPLICANT: STEPHAN, Hans  
; APPLICANT: KREUTZER, Caroline  
; APPLICANT: HERMANN, Thomas  
; APPLICANT: PFEFFERLE, Walter  
; APPLICANT: BINDER, Michael  
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES WHICH CODE FOR THE rpsL GENE  
; FILE REFERENCE: 2042090U0  
; CURRENT APPLICATION NUMBER: US/09/984,711  
; PRIOR FILING DATE: 2001-10-31  
; PRIOR APPLICATION NUMBER: DE10108230.9  
; PRIOR FILING DATE: 2001-02-16  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 5  
; LENGTH: 5096  
; TYPE: DNA  
; ORGANISM: Corynebacterium glutamicum  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (702)..(4196)  
; OTHER INFORMATION:  
US-09-984-711-5

Query Match	38.5%	Score 80;	DB 10;	Length 5096;
Best Local Similarity	66.0%;	Pred. No. 1.3e-12;		
Matches 132;	Conservative	0;	Mismatches 65;	Gaps 1;
QY	9	AACCGCTAGCACTGGCGCGTTCGCCCATACAGTCAACGAAGAGCTGGCGCTGAAC	68	
Db	1578	AACGCGTACGACCTGGCTCGCGTTGGTTCGTACAGATCAACGCGAAGCTGGCGCTTGGT	1637	
QY	69	ACCAATCATCCGATCAACCAGCAGCGTGCACCGAAGAGAGAGCTGCCGCCACCATCGAG	128	
Db	1638	GGCGACCACGATGGTTT---GATGACTCTTACTGAAGAGACATCGCAACCAACCATCGAG	1694	







```

; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Falco, Saverio Carl
; APPLICANT: Fang, Yiwen
; APPLICANT: Hantke, Sabine S.
; APPLICANT: Lee, Jian-Ming
; APPLICANT: Li, Zhongsen
; APPLICANT: Miao, Guo-Hua
; APPLICANT: Morgante, Michele
; APPLICANT: Niu, Xiping
; APPLICANT: Odell, Joan
; APPLICANT: Rafalski, Antoni
; APPLICANT: Sakai, Hajime
; APPLICANT: Zheng, Peizhong
; APPLICANT: Zhu, Qun
; TITLE OF INVENTION: Polynucleotides Encoding Proteins Involved In Plant Metabolism
; FILE REFERENCE: US/10/062,254
; CURRENT APPLICATION NUMBER: 2002-02-01
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: 60/146511
; PRIOR FILING DATE: 1999-07-30
; PRIOR APPLICATION NUMBER: 60/156006
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 60/156899
; PRIOR FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: 60/157287
; PRIOR FILING DATE: 1999-10-01
; PRIOR APPLICATION NUMBER: 60/169767
; PRIOR FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: 60/171054
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: 60/172958
; PRIOR FILING DATE: 1999-12-21
; PRIOR APPLICATION NUMBER: 60/171515
; PRIOR FILING DATE: 1999-12-22
; PRIOR APPLICATION NUMBER: 60/173535
; PRIOR FILING DATE: 1999-12-29
; NUMBER OF SEQ ID NOS: 375
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 111
; LENGTH: 752
; TYPE: DNA
; ORGANISM: Zea mays
US-10-062-254-111

Query Match      18.1%; Score 37.6; DB 12; Length 752;
Best Local Similarity 56.5%; Pred. No. 0.14;
Matches 70; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

Qy  4  AGGAGAGCGCTACGACCTGCGCCGCTGTCGCGCGGATACAAAGGTCAACAAGGTCAACAAGGTGCGGCC 63
Db  236 AGCTGATCGCGCCAGCACCTGCGCCACCGATGCGCGAAGGAGTCCACCAAGAAGTCTCGGCC 295

Qy  64  TCAACACCAATCATCCGATCACCAGCAGCGCTGACCGGAAGAAGAGCTCGTCCGCCACCA 123
Db  296 TCGTCTACCACTCAACATCGCCCCCGGATCGGCGTTGATGAGGAGATCTTCGTGG 355

Qy  124 TCGA 127
Db  356 TCGA 359

RESULT 8
US-09-815-242-4001
; Sequence 4001, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: 2001-03-21
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
```

```

; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4001
; LENGTH: 897
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-815-242-4001

Query Match      18.1%; Score 37.6; DB 10; Length 897;
Best Local Similarity 51.2%; Pred. No. 0.14;
Matches 88; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

Qy  17  CGACCTGGCCCGTGTGCGCGGATACAAAGGTCAACAAGGTGCGCTGCAACCAATCA 76
Db  15  CGCCAGTGGCCGCCCGCCAGCAGCAGGAGACACCTCTCGCGAGCTCCCGCGGAATCA 74

Qy  77  TCGGATCACCACGACGACGCTGACCGAAGAACGACGCTGCGCCACCATCGAGTATCTGGT 136
Db  75  TCCCTATCACACGCGGATGCGCGCAACGAGTACGAGAAGGCCAAGCAGCAGCTGCAGAT 134

Qy  137  CCGCCTGCGCAGGCGCCAGCCAGCATGACCGTCCCGGGCGGGTGGAGGTG 188
Db  135  CGAACTGCTCAAGGTGCGAGCTGGGTGAAGAGAGACCGCGCGCGTGGTG 186

RESULT 9
US-09-815-242-7946
; Sequence 7946, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
```

```
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7946
; LENGTH: 492
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(492)
US-09-815-242-7946
```

```
Query Match          17.8%; Score 37; DB 10; Length 492;
Best Local Similarity 51.5%; Pred. No. 0.19;
Matches 85; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY 21 CTGGCCCGTGTGGCCGATACAAAGCTCAACAAAGAGCTGGGCTGAACACCAATCATCCG 80
    ||| || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 142 CTCACCACCGTGGGCACCATGAGGTGATCGAGCAGATCACCAAGAACCTCAACAAGCTG 201
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 81 ATCACCACGACGCGTGCACCAAGAGACGTGTCGGCCACCACATCGAGTATCTGGTCCGC 140
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 202 ATCGAAGTGTCAAGCTGGTGTCTGTCTCGAAACGCCCATATCGAGCGGAGCTGATG 261
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 141 CTCACGAGCGGCGGACGACGATGACGTCGCCGGCGGGCTCGAG 185
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 262 CTGCTGAAGTCAAGGCCACGCGGCCACGCGCGGCGGAGTCAAG 306
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
RESULT 10
US-09-294-093B-1392
; Sequence 1392, Application US/09294093B
; Patent No. US20010051335A1
; GENERAL INFORMATION:
; APPLICANT: Lalgudi, Raghunath, V.
; APPLICANT: Ito, Laura, Y.
; APPLICANT: Sherman, Bradley, K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN TASSEL
; FILE REFERENCE: PL-0009 US
; CURRENT APPLICATION NUMBER: US/09/294,093B
; PRIOR FILING DATE: 1999-04-16
; PRIOR APPLICATION NUMBER: 60/082,567
; PRIOR FILING DATE: April 21, 1998
; NUMBER OF SEQ ID NOS: 6207
; SOFTWARE: PERL Program
; SEQ ID NO 1392
; LENGTH: 278
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: incyte ID No. US20010051335A1 700344235H1
; NAME/KEY: unsure
; LOCATION: 158
; OTHER INFORMATION: a, t, c, g, or other
US-09-294-093B-1392
```

```
Query Match          17.4%; Score 36.2; DB 10; Length 278;
Best Local Similarity 57.5%; Pred. No. 0.28;
Matches 65; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 4 AGGAGAAGCGCTACGACCTGGCCCGTGTGGCCGATACAAAGGTCAACAAGAGCTGGGCC 63
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 164 AGTGATGCCCGCCAGCTGGCCACCGATGCCGCAAGGAGGTCAACCAAGAGGTGGGCC 223
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 64 TGAACACCAATCATCGCATCACCAACGACGCTGACCAAGAGAGCTGCTC 116
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 224 TCGTCTACCAGTCAACATCGCCCCCAAGAAGATCGCGTTGTATGAGGATC 276
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
RESULT 11
US-09-880-107-3678/c
; Sequence 3678, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 3678
; LENGTH: 2693
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 X04654
US-09-880-107-3678
```

```
Query Match          17.4%; Score 36.2; DB 10; Length 2693;
Best Local Similarity 52.3%; Pred. No. 0.38;
Matches 80; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 39 TACAAGGTCAACAAGAAGCTGGCGCTGAACACCAATCATCCGATCACCACGACGCTG 98
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 273 TACAAGTCAACCAACAAGCTGGGTACGGCTTCCTACCTACCTGCGCGCGTGGCGCTG 214
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 99 ACCGAAGAAGAGCTGCTGCCACCATCGAGTATCTGGTCCGCTGCACGAGGCCAGGCC 158
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 213 TGCCTGGTGGCGGTGACCATGCCCTGGTGTGATGCTGCACGCGCTGCTGCGCTCGGCC 154
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 159 ACGATGACCGCTGCCGGGGGTGCGAGGTGCCG 191
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 153 ACAAGTTCGTGCCCATCAAGGCAAGGCGCG 121
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
RESULT 12
US-09-815-242-7738
; Sequence 7738, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011a
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
```

: PRIOR FILING DATE:	2000-11-27
: PRIOR APPLICATION NUMBER:	60/257,931
: PRIOR FILING DATE:	2000-12-22
: PRIOR APPLICATION NUMBER:	60/269,308
: PRIOR FILING DATE:	2001-02-16
: NUMBER OF SEQ ID NOS:	14110
: SOFTWARE:	FastSeq for Windows version 4.0
: SEQ ID NO 7738	:
: LENGTH:	1437
: TYPE:	DNA
: ORGANISM:	Pseudomonas aeruginosa
: FEATURE:	:
: NAME/KEY:	CDS
: LOCATION:	(1)...(1437)
US-09-815-242-7738	:

  

	Query Match Best Local Similarity Matches	Score 53.6%; Pred. No. 0.5; Conservative	Mismatches 0;	Gaps 0;
--	---	---	------------------	------------

Qy	46	TCAACAAGAAGCTGGCGGTGAACACCATAATCATCCGATCACCAGCAGGCCTTGACCGAAG	105
Db	809	TCCAACAAGTCATCGTGGCGGTGGCCGTGCCCGGTGACCCACCGACCTGCTTTGGCGGCG	868
Qy	106	AAGACTGCTGCCGCCACCATCACGATTATCTGCTCGCGCTGCACGAGGCCACGACGATCA	165
Db	869	ACAGCGCGTGCACCTGGAGGACGCTGGCTTATCTACGTCGACGACCACTCAAAGACCA	928
Qy	166	CCGTGCGCGCGCGGCTCG	183
Db	929	GCCTTCCGGGCGCTTCG	946

  

RESULT 13
US-09-864-761-32324
: Sequence 32324, Application US/09864761
: Patent No. US20020048763Al
: GENERAL INFORMATION:
: APPLICANT: Penn, Sharron G.
: APPLICANT: Rank, David R.
: APPLICANT: Hanzel, David K.
: APPLICANT: Chen, Wensheng
: TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
: FILE REFERENCE: Aemlica-x-1
: CURRENT APPLICATION NUMBER: US/09/864,761
: CURRENT FILING DATE: 2001-05-23
: PRIOR APPLICATION NUMBER: US 60/180,312
: PRIOR FILING DATE: 2000-02-04
: PRIOR APPLICATION NUMBER: US 60/207,456
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: US 09/632,366
: PRIOR FILING DATE: 2000-08-03
: PRIOR APPLICATION NUMBER: GB 24263.6
: PRIOR FILING DATE: 2000-10-04
: PRIOR APPLICATION NUMBER: US 60/236,359
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: PCT/US01/00666
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00667
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00664
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00669
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00665
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00668
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00663
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00662
: PRIOR FILING DATE: 2001-01-30



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

CM nucleic - nucleic search, using sw model  
Run on: November 12, 2002, 16:59:42 ; Search time 1109.13 Seconds  
(without alignments)  
3037.202 Million cell updates/sec

Title: US-09-697-123B-8  
Perfect score: 208  
Sequence: 1 tcaagagagacgctacac.....ccggtgaaacacgacacat 208

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues  
Total number of hits satisfying chosen parameters: 32308132

Minimum DR seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estnu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_estl:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vrt:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_Other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rod:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES				
Result No.	Score	Query Match	Length DR ID	Description
1	160	76.9	1282 9	AI770311 42 Mycoba
2	44	21.2	1019 17	CNS0185F
3	43.2	20.8	630 9	AI405818 GH25953.5
4	43.2	20.8	662 9	AI405617 GH25702.5
5	40.6	19.5	397 12	BF007092 1477141 A
6	40.6	19.5	568 13	BI995346 1031026E0

7	39.2	18.8	516	10	BE056696	BE056696 894011A09
8	39.2	18.8	666	14	BQ817943	BQ817943 1030067C0
9	39.2	18.8	681	14	BQ817941	BQ817941 1030067C0
10	39	18.8	464	13	BM501065	BM501065 PAC000000
11	39	18.8	540	12	BF008145	BF008145 1682830 A
12	39	18.8	710	12	BG321153	BG321153 zmo4_0510
13	39	18.8	762	12	BG836339	BG836339 zmo6_0140
14	39	18.8	996	11	AY104676	AY104676 zea mays
15	39	18.8	1856	11	AY109418	AY109418 zea mays
16	38.8	18.7	656	13	BI960536	BI960536 HVSME002
17	38.6	18.6	701	12	BG836253	BG836253 zmo6_02H0
18	38.4	18.5	413	14	BQ819187	BQ819187 1030076A1
19	38.4	18.5	648	14	BQ579756	BQ579756 WHE2974_L8
20	38.4	18.5	877	13	BI957118	BI957118 HVSME0000
21	38.2	18.4	550	14	BQ767542	BQ767542 EBr008_50
22	38	18.3	511	10	AV628007	AV628007 AV628007
23	38	18.3	543	10	AV631761	AV631761 AV631761
24	38	18.3	554	10	AV631497	AV631497 AV631497
25	38	18.3	566	9	AL825144	AL825144 AL825144
26	38	18.3	725	14	BQ823367	BQ823367 1030110E0
27	37.8	18.2	448	12	BG463919	BG463919 EM1_52_E0
28	37.8	18.2	539	12	BG556957	BG556957 EM1_40_C0
29	37.8	18.2	589	12	BG322881	BG322881 EM1_15_C0
30	37.8	18.2	790	17	BH374681	BH374681 AG-ND-170
31	37.8	18.2	925	17	CNS0091P	AL053013 Drosophil
32	37.6	18.1	431	13	BI097733	BI097733 949017G04
33	37.6	18.1	507	13	BI096671	BI096671 949019B12
34	37.6	18.1	531	13	BM348454	BM348454 MEST291-B
35	37.6	18.1	579	13	BM340507	BM340507 MEST323-F
36	37.6	18.1	586	12	BG842655	BG842655 MEST34-E0
37	37.6	18.1	593	13	BM079854	BM079854 MEST100-E
38	37.6	18.1	597	13	BM333485	BM333485 MEST156-D
39	37.6	18.1	606	13	BM337568	BM337568 MEST209-B
40	37.6	18.1	610	13	BM073467	BM073467 MEST67-H1
41	37.6	18.1	613	13	BM340956	BM340956 MEST328-C
42	37.6	18.1	622	13	BM337575	BM337575 MEST209-C
43	37.6	18.1	651	13	BM332944	BM332944 MEST181-G
44	37.6	18.1	659	13	BM266851	BM266851 MEST386-H
45	37.6	18.1	659	13	BM337566	BM337566 MEST209-H

ALIGNMENTS

RESULT 1	AI770311/c	AI770311	1282 bp	mRNA	linear	EST 24-JAN-2000
LOCUS	42 Mycobacterium anaerobic stationary phase library Mycobacterium smegmatis CDNA, mRNA sequence.	AI770311				
DEFINITION	42 Mycobacterium anaerobic stationary phase library Mycobacterium smegmatis CDNA, mRNA sequence.	AI770311				
ACCESSION	AI770311	AI770311.1	GI:6742680			
VERSION	EST.					
KEYWORDS	Mycobacterium smegmatis.					
SOURCE	Mycobacterium smegmatis					
ORGANISM	Bacteria: Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.					
REFERENCE	1 (bases 1 to 1282)					
AUTHORS	Murugasu-Oei, B., Tay, A. and Dick, T.					
TITLE	Upregulation of stress response genes and ABC transporters in anaerobic stationary-phase Mycobacterium smegmatis					
JOURNAL	Mol. Gen. Genet. 262 (4-5), 677-682 (1999)					
MEDLINE	20092472					
COMMENT	Contact: Murugasu-Oei, B. Mycobacterium Laboratory Institute of Molecular and Cell Biology 30 Medical Drive, Singapore 117609, Republic of Singapore Tel: 65 874 3011 Fax: 65 779 1117 Email: mcbom@imcb.nus.edu.sg Insert Length: 1282 Std Error: 0.00 Seq primer: T3 Forward: T7 Backward.					
FEATURES	Location/Qualifiers					
source	1. .1282					



```
Query Match      20.8%; Score 43.2; DB 9; Length 630;
Best Local Similarity 51.6%; Pred. No. 0.99;
Matches 99; Conservative 0; Mismatches 93; Indels 0; Gaps 0;

QY 9 AAGCGCTACGACCTGCCCGTGTGCGCGATACAAAGGTCAACAAGAGCTGGCGCTGAAC 68
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 190 AAGCAAAAGATCGTGCGACATTTTGAAGCGGTTCAAAGTGTGAGCGAGGAGCTGGAGGAG 249
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 69 ACCAATCATCCAGTACCACGACGACGCTGACCGAAGAGAGCTGCTGCCACCATCCAG 128
    || ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 250 TCCGGTCATCGAGGACGCGCGCTGCGGAGAACCGGACCTGGATGGCGCTTTCAG 309
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 129 TATCTGGTCCGCTGACGAGCGCAGGACACGATGACCTGCGCGGGCGGGTCCAGGTG 188
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 310 GAGCTGGAGTCCGTCTCTGTTGCGAGGCGACGACTCCGGACCGGACATGGACAGCATA 369
    || ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 189 CCGGTGGAAACC 200
    ||||| | |
Db 370 TCGGTGGGCTCC 381

RESULT 5
BF007092      397 bp      mRNA      linear      EST 06-OCT-2000
LOCUS      1471141 Amblyomma americanum adult Lambda zap Express Amblyomma
DEFINITION      americanum cDNA, mRNA sequence.
ACCESSION      BF007092
VERSION      BF007092.1 GI:10707367
KEYWORDS      EST.
SOURCE      Amblyomma americanum.
ORGANISM      Amblyomma americanum
Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
Parasitiformes; Ixodida; Ixodidae; Amblyomma.
REFERENCE      1 (bases 1 to 397)
AUTHORS      Hill,C.A and Gutierrez,J.A.
TITLE      Analysis of the expressed genome of the lone star tick, Amblyomma
JOURNAL      americanum (Acari:Ixodidae) using an expressed sequence tag
COMMENT      approach
Microb. Comp. Genomics 5 (2), 89-101 (2000) In press
Contact: Hill CA
Animal Science Discovery Research
Elanco Animal Health, A Division of Eli Lilly and Company
PO Box 708, 2001 West Main St., Greenfield, IN 46140, USA
Tel: 317 277 0826
Fax: 317 277 4522
Email: HILL_CATHERINE_A@LILLY.COM.

FEATURES
source
1..397
/organism="Amblyomma americanum"
/db_xref="taxon:6943"
/clone_lib="Amblyomma americanum adult Lambda zap Express"
/sex="Male, Female"
/dev_stage="Adult"
/note="Vector: pBK-CMV; Site_1: EcoRI; Site_2: XhoI"

BASE COUNT      124 a 110 c 124 g 36 t 3 others
ORIGIN

Query Match      19.5%; Score 40.6; DB 12; Length 397;
Best Local Similarity 55.2%; Pred. No. 4;
Matches 79; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 35 CCGATACAGGTCAACAAGAGCTGGCGCTGAACACCAATCATCCGATCACCACGACGAC 94
    || ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 223 CCAGGACGAGCTCATCAACAGCTCAACAAGGAGAAGAAGAGTTCGAGGAGCAGAACCA 282
    || ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 95 GCTGACCGAAGAGAGCTGCTGCCACCATCGAGTATCTGTGTCGCCCTGCCAGGGCCA 154
    || ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 283 GAAGACCGCCGAAGACCTCCAGGCCACCGAGGACAAAGGTGAACACCACTGAACAAAGTCAA 342
    || ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 155 GGCCACGATGACCGTCCCGGGCG 177
    ||||| | | | | |
Db 343 GCCCAAGCTGGAGACAGCGCTCG 365

RESULT 6
BI995346      568 bp      mRNA      linear      EST 25-OCT-2001
LOCUS      1031026E07.y2 C. reinhardtii CC-1690, Stress II (normalized),
DEFINITION      Lambda zap II Chlamydomonas reinhardtii cDNA, mRNA sequence.
ACCESSION      BI995346
VERSION      BI995346.1 GI:16430136
KEYWORDS      EST.
```







```

BASE COUNT      78 a   137 c   181 g   67 t   1 others
ORIGIN

Query Match      18.8%; Score 39; DB 13; Length 464;
Best Local Similarity 52.8%; Pred. No. 10;
Matches 84; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 47 CAACAAGAGCTGGCGCTGAACACGATCATCCGATCACCACGACGCGTGAACCGAAGA 106
    ||| | ||| ||| ||| ||| | ||| | ||| | ||| | ||| | ||| | ||| |
Db 129 CAAGAGCGCTCTCTCTGGGTGCAACCGGTGCACAAGAGCGTGGCGTCTCTCCA 188
    ||| | ||| ||| ||| ||| | ||| | ||| | ||| | ||| | ||| | ||| |
QY 107 AGACGCTGCTGCCACCATCGAGTAGTCTGGTCCGCTGCACAGAGGCGCAGCATGAC 166
    ||| | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 189 GCGCGCGCGCCCAATAACAGTAGTCCAGGTCATCTCGAGTATACATCGACGAGGTGGGC 248
    ||| | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 167 GCTGCCGCGCGGGTCGAGGTGCGGTTGGAACCGAGCA 205
    ||| | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 249 GGGACGCGCGGACGCTGCTCGCGCGGACCCCTACGA 287
    ||| | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 11
BF008145      540 bp      mRNA      linear      EST 06-OCT-2000
LOCUS      1682830 Amblyomma americanum adult Lambda Zap Express Amblyomma
DEFINITION      americanum cDNA. mRNA sequence.
ACCESSION      BF008145
VERSION      BF008145.1 GI:10708420
KEYWORDS      EST.
SOURCE      Amblyomma americanum.
ORGANISM      Amblyomma americanum.
Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
Parasitiformes; Ixodida; Ixodidae; Amblyomma.
REFERENCE:
AUTHORS      Hill,C.A and Gutierrez,J.A.
TITLE      Analysis of the expressed genome of the lone star tick, Amblyomma
americanum (Acari:Ixodidae) using an expressed sequence tag
approach
JOURNAL      Microb. Comp. Genomics 5 (2), 89-101 (2000) In press
COMMENT      Contact: Hill CA
Animal Science Discovery Research
Elanco Animal Health, A Division of Eli Lilly and Company
PO Box 708, 2001 West Main St., Greenfield, IN 46140, USA
Tel: 317 277 0826
Fax: 317 277 4522
Email: HILL_CATHERINE_A@LILLY.COM.
FEATURES
source
    1..540
        /organism="Amblyomma americanum"
        /db_xref="taxon:6943"
        /clone_lib="Amblyomma americanum adult Lambda Zap Express"
        /sex="Male, Female"
        /dev_stage="Adult"
        /note="Vector: pBK-CMV; Site_1: EcoRI; Site_2: XhoI"

BASE COUNT      164 a   149 c   175 g   52 t
ORIGIN

Query Match      18.8%; Score 39; DB 12; Length 540;
Best Local Similarity 54.5%; Pred. No. 10;
Matches 78; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

QY 35 CCGATACAAGTCAACAAGAGCTGGCGCTGAACACCAATCATCCGATCACACGAGCAG 94
    ||| | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 362 CCAGGACGAGTCAACAAGAGCTCAACAAGAGAGAGACGATTCGAGGAGCAACCA 421
    ||| | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 95 GCTGACCGAAGACAGCTCTCGCCACCATCGAGTATCTGCTCGGCTGCACGAGGCCA 154
    ||| | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 422 GAAGACCGCGAGACCTCCAGCAACCGAGAGAGAGGTGAACCACTGAACAAGGTCAA 481
    ||| | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 155 GGCACCATGACCGTGGCGGCG 177
    ||| | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 482 GGCCAAGCTGGACAGACGCTCG 504
    ||| | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 12
BG321153
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

```

```

BG321153      710 bp      mRNA      linear      EST 27-FEB-2001
cDNA clone Zm04_AAFc_ECORC_cold_stressed_maize_seedlings Zea mays
BG321153
BG321153.1 GI:13150831
EST.
Zea mays.
Zea mays.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 710)
Singh,J.A., Wakui,K., Couroux,P., De Moors,A., Harris,L.J., Hattori
,J.I., Ouellet,T., Robert,L.S., Sprott,D. and Tinker,N.A.
Expressed Sequence Tags from Cold-Stressed Maize Seedlings
Unpublished (2001)
Contact: Singh,J.A.
Eastern Cereal and Oilseed Research Centre
Agriculture and Agri-food Canada
KW Neatby Bldg., Central Experimental Farm, Ottawa, Ontario, K1A
0C6, Canada
Tel: (613) 759-1662
Fax: (613) 759-1701
Email: singhja@em.agr.ca.
Location/Qualifiers
    1..710
        /organism="Zea mays"
        /cultivar="CO328"
        /db_xref="taxon:4577"
        /clone_lib="Zm04_05f06"
        /tissue_type="Leaf, crown"
        /note="Vector: Bluescript SK-/XhoI-EcoRI; Site_1: Eco RI;
        Site_2: Xho I; Lower temperature 50 C / hour from 22 to
        120C; bring to 50 in 1 hour from 120C. Leave at 50C 2 days
        photoperiod 16 hours. Light intensity was 125 uE-l.
        Library prepared by in vivo mass excision from amplified
        library."

```

```

FEATURES
source
    1..710

```

```

BASE COUNT      118 a   263 c   213 g   80 t   36 others
ORIGIN

Query Match      18.8%; Score 39; DB 12; Length 710;
Best Local Similarity 59.5%; Pred. No. 10;
Matches 66; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

QY 45 GTCAACAAGAAGCTGGCGCTGAACACCAATCATCCGATCACCACGACGCGTGACCGAA 104
    ||| | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 388 GGCATCCAGGAMCTGGCCAGGTGCCCATGAGCAGATCGACGCCCGCAAGCTCACCAAG 447
    ||| | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 105 GAAGAGCTGTCGCCACCATCGAGTATCTGTCGCGCTGCACGAGGGCCAG 155
    ||| | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 448 GAGCAGCTCGACGCTGCGCTCAAGTTGCTCGGCTCTTCACGCGCGCAAG 498
    ||| | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

FEATURES
source
    1..710

```

```

BASE COUNT      118 a   263 c   213 g   80 t   36 others
ORIGIN

Query Match      18.8%; Score 39; DB 12; Length 710;
Best Local Similarity 59.5%; Pred. No. 10;
Matches 66; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

QY 45 GTCAACAAGAAGCTGGCGCTGAACACCAATCATCCGATCACCACGACGCGTGACCGAA 104
    ||| | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 388 GGCATCCAGGAMCTGGCCAGGTGCCCATGAGCAGATCGACGCCCGCAAGCTCACCAAG 447
    ||| | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 105 GAAGAGCTGTCGCCACCATCGAGTATCTGTCGCGCTGCACGAGGGCCAG 155
    ||| | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 448 GAGCAGCTCGACGCTGCGCTCAAGTTGCTCGGCTCTTCACGCGCGCAAG 498
    ||| | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 13
BG836339
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

```

```

BG836339      762 bp      mRNA      linear      EST 25-MAY-2001
Zm06_01a02_R
Zm06_AAFc_ECORC_Fusarium-graminearum_inoculated_corn_eartip Zea
mays cDNA clone Zm06_01a02, mRNA sequence.
BG836339
BG836339.1 GI:14202662
EST.
Zea mays.
Zea mays.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 762)
Harris,L.J., Balcerzak,M., Allard,S., Saparno,A., Couroux,P., De
Moors,A., Hattori,J.I., Ouellet,T., Robert,L.S., Singh,J.A, Sprott

```

.D. and Tinker N.A.  
Expressed Sequence Tags from Maize Ear Tips 48 Hours after Silk  
Channel Inoculation with Fusarium graminearum

TITLE  
JOURNAL  
COMMENT  
Unpublished (2001)  
Contact: Harris, Linda J.  
Eastern Cereal and Oilseed Research Centre  
Agriculture and Agri-Food Canada  
Bldg. 21, Central Experimental Farm, Ottawa, Ontario, K1A 0C6,  
CANADA  
Tel: (613) 759-1314  
Fax: (613) 759-6566  
Email: harrislj@em.agr.ca.

FEATURES  
source  
location/Qualifiers  
1..762  
/organism="Zea mays"  
/cultivar="CO387"  
/db\_xref="taxon:4577"  
/clone="zn06\_01a02"  
/clone\_lib="zn06\_AAPF\_ECOHC\_Fusarium\_graminearum\_inoculate  
d\_corn\_eartip"  
/tissue\_type="2.5 cm of top of unfertilized ear"  
/dev\_stage="6-7 days post-silk emergence"  
/note="Vector: Bluescript SK+/XhoI-EcoRI; Site\_1: EcoRI;  
Site\_2: XhoI; Field-grown unpollinated maize ears were  
silk channel-inoculated in the morning (~10 am) with 1 ml  
of a Fusarium graminearum macroconidial suspension (500  
000 spores/ml) and whole ears were collected and  
immediately frozen in liquid nitrogen 48 hours later."  
BASE COUNT 124 a 223 c 280 g 123 t 12 others  
ORIGIN

Query Match	18.8%	Score 39;	DB 12;	Length 762;	
Best Local Similarity	52.8%;	Pred. No. 11;			
Matches 84;	Conservative	0;	Mismatches 75;	Indels 0;	Gaps 0;

```

QY 47 CAACAAGAGCTGGCGCTGAACACCATCATCCCATCACGACGACGCTGACCGGAAGA 106
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 190 CAAGAGCGCGTCTCTCTGGGTCAACCCGGTGCAACGAGCGTGCCTGCCCTCCA 249
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 107 AGAGTCGTGCCACCATCGAGTAGTCTGTCGGCTGCAGGAGGGCCAGGCACGATGAC 166
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 250 CGCCGGCGCGCCATAAACGAGTCCCAGGTCATCTCGACTACATCAGCAGGTGTGGC 309
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 167 CTTCCCGCGGGTTCGAGGTGCGCGTGGAAACCGACGA 205
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 310 GGGACGGGGCGGACGCTGTCGGGACGACCCCTACGA 348
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

RESULT 14  
AY104676

LOCUS  
AY104676  
DEFINITION Zea mays PC0085803 mRNA sequence.  
ACCESSION AY104676  
VERSION AY104676.1 GI:21207754  
KEYWORDS HTC.  
SOURCE Zea mays.  
ORGANISM Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC  
clade; Panicoideae; Andropogoneae; Zea.  
1 (bases 1 to 996)  
Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,  
Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.  
Maize Mapping Project/DuPont Consensus Sequences for Design of  
Overgo Probes  
Unpublished (2002)  
REFERENCE 2 (bases 1 to 996)  
AUTHORS Coe,E.C.  
TITLE Direct Submission  
JOURNAL Submitted (25-APR-2002) Maize Mapping Project, University of  
Missouri, Columbia, MO 65211, USA  
LOCATION/Qualifiers  
1..996

/organism="Zea mays" /db_xref="MaizeDB:634819" /db_xref="taxon:4577" /clone_lib="Maize Mapping Project/DuPont Cornsensus Library" /note="this sequence is part of a project of EST assemblies resulting from the application of public contigs to seed DuPont contigs; this resource was assembled by DuPont as part of a collaboration for the overgo addressing of BACs in conjunction with the Maize Mapping Project" BASE COUNT 193 a 286 c 326 g 191 t ORIGIN									
Query Match 18.8% Score 39; DB 11; Length 996; Best Local Similarity 52.8%; Pred.No.11; Matches 84; Conservative 0; Mismatches 75; Indels 0; Caps 0;									
Qy	47	CAACAAGAGCTGGCGCTGAACACCAATCATCCGATCACCACGACGCGTGACCCGAAGA	106						
Db	207	CAAGAGCGCTCTCTCTGGGATCCAAACCGCGTGCAACAAGAGCGTGCGGTCTCTCTCCA	266						
Qy	107	AGAGCTCGTCCCGCACCATCGAGTATCTGTGTCGCTCGACGAGGCGCAGGCCAGGATGAC	166						
Db	267	CGCGCGCGCGCCATCAACAGAGTCCCGAGGTCTCTCGAGTACATCGACGAGGTGTGGGC	326						
Qy	167	CGTCCGCGCGGGTCTGAGGTGCGCGTGGAAACGACGA	205						
Db	327	GGGACGCGCGGACGTGTCGCCGCGGACCCCTACGA	365						
RESULT 15 AY109418 LOCUS AY109418 1856 bp mRNA linear HTC 25-MAY-2002 DEFINITION Zea mays CL2386_1 mRNA sequence. ACCESSION AY109418 VERSION AY109418.1 GI:21213135 KEYWORDS HTC. SOURCE Zea mays. ORGANISM Zea mays									
REFERENCE AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea. 1 (bases 1 to 1856) Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S., Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V. Maize Mapping Project/DuPont Consensus Sequences for Design of Overgo Probes Unpublished (2002) 2 (bases 1 to 1856) Coe,E.C. Direct Submission Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA Location/Qualifiers 1. .1856 /organism="Zea mays" /db_xref="MaizeDB:630950" /db_xref="taxon:4577" /clone_lib="Maize Mapping Project/DuPont Cornsensus Library" /note="this sequence is part of a project of EST assemblies resulting from the application of public contigs to seed DuPont contigs; this resource was assembled by DuPont as part of a collaboration for the overgo addressing of BACs in conjunction with the Maize Mapping Project" BASE COUNT 320 a 548 c 585 g 281 t 122 others ORIGIN									
Query Match 18.8% Score 39; DB 11; Length 1856;									

Best Local Similarity 59.5%; Pred. No. 12;  
Matches 66; Conservative 0; Mismatches 45; Indels 0; Gaps 0;  
QY 45 GTCAACAGAGCTGGGCTGAACACCAATCATTCGATCACCCAGCAGCGCTGACCGAA 104  
Dh 140 GGCATCCAGGACCTGGCCAGGCTCCCCATCGAGCAGATCGAGCGCGCCCAAGCTCACCAG 199  
QY 105 GAAGAGCTGCTGCCACCATCGAGTATCTGGTCCGCCCTGCACGAGGGCCAG 155  
Db 200 CACGAGCTCGAGCTCCGGCTCAAGTGGCTCGGCCCTCTTCCACCGCGCGCAAG 250

Search completed: November 13, 2002, 03:50:54  
Job time : 1122.63 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

QM nucleic - nucleic search, using sw model  
Run on: November 12, 2002, 16:50:46 ; Search time 681.89 Seconds  
(without alignments)  
9517.553 Million cell updates/sec

Title: US-09-697-123b-13  
Perferr score: 223  
Sequence: 1 tcaaggagaagcgctacgac.....ccgggtggaacacgacgacat 223

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :	
GenEmbl.*	
1: gb_ba.*	
2: gb_htg.*	
3: gb_in.*	
4: gb_om.*	
5: gb_ov.*	
6: gb_pat.*	
7: gb_ph.*	
8: gb_pl.*	
9: gb_pr.*	
10: gb_ro.*	
11: gb_sts.*	
12: gb_sy.*	
13: gb_un.*	
14: gb_vi.*	
15: em_ba.*	
16: em_fun.*	
17: em_hum.*	
18: em_in.*	
19: em_mu.*	
20: em_om.*	
21: em_or.*	
22: em_ov.*	
23: em_pat.*	
24: em_ph.*	
25: em_pl.*	
26: em_ro.*	
27: em_sts.*	
28: em_un.*	
29: em_vi.*	
30: em_htg_hum.*	
31: em_htg_inv.*	
32: em_htg_other.*	
33: em_htg_mus.*	
34: em_htg_pln.*	
35: em_htg_rod.*	
36: em_htg_mam.*	
37: em_htg_vrt.*	
38: em_sy.*	
39: em_htgo_hum.*	
40: em_htgo_mus.*	
41: em_htgo_other.*	

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	123.8	55.5	610	1	MTU318818	AJ318818 Mycobacte
2	123.8	55.5	610	1	MTU318819	AJ318819 Mycobacte
3	123.8	55.5	616	1	MTU318813	AJ318813 Mycobacte
4	123.8	55.5	618	1	MTU318815	AJ318815 Mycobacte
5	123.8	55.5	618	1	MTU318817	AJ318817 Mycobacte
6	123.8	55.5	633	1	MTU318814	AJ318814 Mycobacte
7	123.8	55.5	637	1	MTU318816	AJ318816 Mycobacte
8	123.8	55.5	639	1	MTU318821	AJ318821 Mycobacte
9	123.8	55.5	970	6	I50706	Sequence 1
10	123.8	55.5	3534	6	AX111339	Sequence 1
11	123.8	55.5	3853	1	MTU12205	U12205 Mycobacteri
12	123.8	55.5	5084	1	MSGRPOB	L27989 Mycobacteri
13	123.8	55.5	19352	1	AE006964	AE006964 Mycobacte
14	123.8	55.5	19770	1	MTC1376	295972 Mycobacteri
15	122.8	55.1	3752	1	MSU24494	U24494 Mycobacteri
16	111	49.8	3447	6	AR067447	AR067447 Sequence
17	111	49.8	37617	1	MLB1790G	Z14314 M. leprae ge
18	111	49.8	348950	1	MLEPRTN7	AL583923 Mycobacte
19	88.4	39.6	3941	1	AF242549	AF242549 Amycolato
20	68.8	30.9	32923	1	SGD82	AL160431 Streptomy
21	62.8	28.2	3495	6	AX120631	Sequence
22	62.8	28.2	328050	1	AP005275	AP005275 Corynebac
23	62.8	28.2	349980	6	AX127144	Sequence
24	46.8	21.0	204050	1	AL646070	AL646070 Ralstonia
25	45.2	20.3	121544	2	AC092779	AC092779 Oryza sat
26	45.2	20.3	150446	2	AC092779	AC092779 Oryza sat
27	44.6	20.0	139467	1	D90905	D90905 Synecocyst
28	44.2	19.8	204050	1	AL646073	AL646073 Ralstonia
29	42.2	18.9	2164	1	SC2C3	AL391157 Streptomy
30	42	18.8	3849	6	AX058889	Sequence
31	42	18.8	10809	1	AE002027	AE002027 Deinococc
32	42	18.8	12744	1	AF204401	AF204401 Streptomy
33	41.4	18.6	2771	1	PAU70470	U70470 Pseudomonas
34	41.4	18.6	12948	1	AE004589	AE004589 Pseudomon
35	41.4	18.6	188050	1	AL646072	AL646072 Ralstonia
36	41.2	18.5	3304	6	E03373	E03373 DNA sequenc
37	41.2	18.5	3306	6	I28835	I28835 Sequence 71
38	41	18.4	149777	2	AP003573	AP003573 Oryza sat
39	40.8	18.3	496	1	SBU72186	U72186 Streptomyce
40	40.8	18.3	599	1	AF130448	AF130448 Deinococc
41	40.8	18.3	7954	1	TAQ19223	Y19223 Thermus aqu
42	40.8	18.3	166753	8	AP002866	AP002866 Oryza sat
43	40.6	18.2	1057	1	AB085278	AB085278 Unculture
44	40.6	18.2	1057	1	AB085282	AB085282 Unculture
45	40.6	18.2	39207	1	SC4G1	AL391039 Streptomy

ALIGNMENTS

RESULT 1	MTU318818	610 bp	DNA	circular BCT 09-AUG-2002
LOCUS	MTU318818			
DEFINITION	Mycobacterium tuberculosis partial rpoB gene for RNA polymerase			
ACCESSION	AJ318818			
VERSION	AJ318818.1 GI:22208412			
KEYWORDS	RNA polymerase beta subunit; rpoB gene.			
SOURCE	Mycobacterium tuberculosis.			
ORGANISM	Mycobacterium tuberculosis.			
REFERENCE	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;			
AUTHORS	Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium tuberculosis complex.			
	Herrera, L., Jimenez, M.S. and Saez, J.A.			

```

TITLE      Molecular analysis of rifampin-resistant Mycobacterium tuberculosis
JOURNAL    isolated in Spain (1996-2001). Description of new alleles into rpoB
AUTHORS    gene and review
TITLE      2 (bases 1 to 610)
JOURNAL    Herrera, L.
FEATURES   Direct Submission
SOURCE     Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro
           Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5,
           Majadahonda. Madrid. 28220, SPAIN
           Location/Qualifiers
           1..610
           /organism="Mycobacterium tuberculosis"
           /isolate="1415-97"
           /db_xref="taxon:1773"
           1..610
           /gene="rpoB"
           <1..>610
           /codon_start=1
           /product="RNA polymerase beta subunit"
           /protein_id="CAC87035.1"
           /db_xref="GI:22208413"
           /translation="LDIYKLRPGEPPTKESAOQLLENLFKERYDLARVGRYKVNK
           KLGLHVGEPITSSLTTEEDVVATIEYLRLHEGQTTMTVPGVEVPVETDDIDHFGNR
           RLRTVGEIIQNOIRVGMSEMRVVRMTTQDVEAITPOTLINIRPVVAALKEFFGTS
           QPSQFMGQNNPLSLGTHKRRLSALPGGLSRERAGLEVRDVHP"
BASE COUNT 122 a 191 c 202 g 95 t
ORIGIN
Query Match      55.5%; Score 123.8; DB 1; Length 610;
Best Local Similarity 76.7%; Pred. No. 8.7e-14;
Matches 171; Conservative 0; Mismatches 37; Indels 15; Gaps 1;
OY 1 TCAAGGAGAGCGCTACGACCTGGCCGCGCTGCGCGCTACAAAGTCAACAAGAGCTCG 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 80 TCAAGGAGAGAGCGCTACGACCTGGCCGCGCTGCGCGCTACAAAGTCAACAAGAGCTCG 139
OY 61 GCCTGAACACCGATCATCCGATCACCACGACGCTGACCGCAAGACGCTCGTCGCCA 120
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 140 GCCTGCATGTTCGGCGAGCCCATCAGCTGCTCGACGCTGACCGCAAGACGCTCGTCGCCA 199
OY 121 CCATCGAGTACCTGGTTCGCTCGACGACGCTCTCAGGCTGGCCAGGCCCGCTTATGA 180
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 200 CCATCGAATATCTGCTCCGCTTCACGAGG-----GTCAGACCACGATGA 244
OY 181 CTGTCCCGCGCGGTCGAGGTGCGCGTGGAAACCGACGACAT 223
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 245 CCGTTCCGCGCGGTCGAGGTGCGCGTGGAAACCGACGACAT 287

RESULT 2
MTU318819      610 bp DNA circular BCT 09-AUG-2002
LOCUS          Mycobacterium tuberculosis partial rpoB gene for RNA polymerase
DEFINITION    beta subunit, isolate 1417-97.
ACCESSION     AJ318819
VERSION       AJ318819.1 GI:22208414
KEYWORDS      RNA polymerase beta subunit; rpoB gene.
SOURCE        Mycobacterium tuberculosis.
ORGANISM      Mycobacterium tuberculosis.
               Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
               Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
               tuberculosis complex.
REFERENCE     1
AUTHORS       Herrera, L., Jimenez, M.S. and Saez, J.A.
TITLE         Molecular analysis of rifampin-resistant Mycobacterium tuberculosis
               isolated in Spain (1996-2001). Description of new alleles into rpoB
               gene and review
JOURNAL       unpublished
REFERENCE     2 (bases 1 to 610)
AUTHORS       Herrera, L.
TITLE         Direct Submission
JOURNAL       Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro
               Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5,
               Majadahonda. Madrid. 28220, SPAIN
               Location/Qualifiers
               1..616
               /organism="Mycobacterium tuberculosis"
               /isolate="1415-97"
               /db_xref="taxon:1773"
               1..610
               /gene="rpoB"
               <1..>610
               /codon_start=1
               /product="RNA polymerase beta subunit"
               /protein_id="CAC87035.1"
               /db_xref="GI:22208413"
               /translation="LDIYKLRPGEPPTKESAOQLLENLFKERYDLARVGRYKVNK
               KLGLHVGEPITSSLTTEEDVVATIEYLRLHEGQTTMTVPGVEVPVETDDIDHFGNR
               RLRTVGEIIQNOIRVGMSEMRVVRMTTQDVEAITPOTLINIRPVVAALKEFFGTS
               QPSQFMGQNNPLSLGTHKRRLSALPGGLSRERAGLEVRDVHP"
BASE COUNT 122 a 191 c 202 g 95 t
ORIGIN
Query Match      55.5%; Score 123.8; DB 1; Length 610;
Best Local Similarity 76.7%; Pred. No. 8.7e-14;
Matches 171; Conservative 0; Mismatches 37; Indels 15; Gaps 1;
OY 1 TCAAGGAGAGCGCTACGACCTGGCCGCGCTGCGCGCTACAAAGTCAACAAGAGCTCG 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 80 TCAAGGAGAGAGCGCTACGACCTGGCCGCGCTGCGCGCTACAAAGTCAACAAGAGCTCG 139
OY 61 GCCTGAACACCGATCATCCGATCACCACGACGCTGACCGCAAGACGCTCGTCGCCA 120
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 140 GCCTGCATGTTCGGCGAGCCCATCAGCTGCTCGACGCTGACCGCAAGACGCTCGTCGCCA 199
OY 121 CCATCGAGTACCTGGTTCGCTCGACGACGCTCTCAGGCTGGCCAGGCCCGCTTATGA 180
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 200 CCATCGAATATCTGCTCCGCTTCACGAGG-----GTCAGACCACGATGA 244
OY 181 CTGTCCCGCGCGGTCGAGGTGCGCGTGGAAACCGACGACAT 223
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 245 CCGTTCCGCGCGGTCGAGGTGCGCGTGGAAACCGACGACAT 287

RESULT 2
MTU318819      610 bp DNA circular BCT 09-AUG-2002
LOCUS          Mycobacterium tuberculosis partial rpoB gene for RNA polymerase
DEFINITION    beta subunit, isolate 1417-97.
ACCESSION     AJ318819
VERSION       AJ318819.1 GI:22208414
KEYWORDS      RNA polymerase beta subunit; rpoB gene.
SOURCE        Mycobacterium tuberculosis.
ORGANISM      Mycobacterium tuberculosis.
               Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
               Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
               tuberculosis complex.
REFERENCE     1
AUTHORS       Herrera, L., Jimenez, M.S. and Saez, J.A.
TITLE         Molecular analysis of rifampin-resistant Mycobacterium tuberculosis
               isolated in Spain (1996-2001). Description of new alleles into rpoB
               gene and review
JOURNAL       unpublished
REFERENCE     2 (bases 1 to 610)
AUTHORS       Herrera, L.
TITLE         Direct Submission
JOURNAL       Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro
               Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5,
               Majadahonda. Madrid. 28220, SPAIN
               Location/Qualifiers
               1..616
               /organism="Mycobacterium tuberculosis"
               /isolate="1415-97"
               /db_xref="taxon:1773"
               1..610
               /gene="rpoB"
               <1..>610
               /codon_start=1
               /product="RNA polymerase beta subunit"
               /protein_id="CAC87035.1"
               /db_xref="GI:22208413"
               /translation="LDIYKLRPGEPPTKESAOQLLENLFKERYDLARVGRYKVNK
               KLGLHVGEPITSSLTTEEDVVATIEYLRLHEGQTTMTVPGVEVPVETDDIDHFGNR
               RLRTVGEIIQNOIRVGMSEMRVVRMTTQDVEAITPOTLINIRPVVAALKEFFGTS
               QPSQFMGQNNPLSLGTHKRRLSALPGGLSRERAGLEVRDVHP"
BASE COUNT 122 a 191 c 202 g 95 t
ORIGIN
Query Match      55.5%; Score 123.8; DB 1; Length 610;
Best Local Similarity 76.7%; Pred. No. 8.7e-14;
Matches 171; Conservative 0; Mismatches 37; Indels 15; Gaps 1;
OY 1 TCAAGGAGAGCGCTACGACCTGGCCGCGCTGCGCGCTACAAAGTCAACAAGAGCTCG 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 80 TCAAGGAGAGAGCGCTACGACCTGGCCGCGCTGCGCGCTACAAAGTCAACAAGAGCTCG 139
OY 61 GCCTGAACACCGATCATCCGATCACCACGACGCTGACCGCAAGACGCTCGTCGCCA 120
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 140 GCCTGCATGTTCGGCGAGCCCATCAGCTGCTCGACGCTGACCGCAAGACGCTCGTCGCCA 199
OY 121 CCATCGAGTACCTGGTTCGCTCGACGACGCTCTCAGGCTGGCCAGGCCCGCTTATGA 180
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 200 CCATCGAATATCTGCTCCGCTTCACGAGG-----GTCAGACCACGATGA 244
OY 181 CTGTCCCGCGCGGTCGAGGTGCGCGTGGAAACCGACGACAT 223
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 245 CCGTTCCGCGCGGTCGAGGTGCGCGTGGAAACCGACGACAT 287

RESULT 3
MTU318813      616 bp DNA circular BCT 09-AUG-2002
LOCUS          Mycobacterium tuberculosis partial rpoB gene for RNA polymerase
DEFINITION    beta subunit, isolate 1763-97.
ACCESSION     AJ318813
VERSION       AJ318813.1 GI:22208402
KEYWORDS      RNA polymerase beta subunit; rpoB gene.
SOURCE        Mycobacterium tuberculosis.
ORGANISM      Mycobacterium tuberculosis.
               Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
               Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
               tuberculosis complex.
REFERENCE     1
AUTHORS       Herrera, L., Jimenez, M.S. and Saez, J.A.
TITLE         Molecular analysis of rifampin-resistant Mycobacterium tuberculosis
               isolated in Spain (1996-2001). Description of new alleles into rpoB
               gene and review
JOURNAL       unpublished
REFERENCE     2 (bases 1 to 616)
AUTHORS       Herrera, L.
TITLE         Direct Submission
JOURNAL       Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro
               Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5,
               Majadahonda. Madrid. 28220, SPAIN
               Location/Qualifiers
               1..616
               /organism="Mycobacterium tuberculosis"
               /isolate="1415-97"
               /db_xref="taxon:1773"
               1..610
               /gene="rpoB"
               <1..>610
               /codon_start=1
               /product="RNA polymerase beta subunit"
               /protein_id="CAC87036.1"
               /db_xref="GI:22208415"
               /translation="LDIYKLRPGEPPTKESAOQLLENLFKERYDLARVGRYKVNK
               KLGLHVGEPITSSLTTEEDVVATIEYLRLHEGQTTMTVPGVEVPVETDDIDHFGNR
               RLRTVGEIIQNOIRVGMSEMRVVRMTTQDVEAITPOTLINIRPVVAALKEFFGTS
               QPSQFMGQNNPLSLGTHKRRLSALPGGLSRERAGLEVRDVHP"
BASE COUNT 122 a 191 c 202 g 95 t
ORIGIN
Query Match      55.5%; Score 123.8; DB 1; Length 610;
Best Local Similarity 76.7%; Pred. No. 8.7e-14;
Matches 171; Conservative 0; Mismatches 37; Indels 15; Gaps 1;
OY 1 TCAAGGAGAGCGCTACGACCTGGCCGCGCTGCGCGCTACAAAGTCAACAAGAGCTCG 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 80 TCAAGGAGAGAGCGCTACGACCTGGCCGCGCTGCGCGCTACAAAGTCAACAAGAGCTCG 139
OY 61 GCCTGAACACCGATCATCCGATCACCACGACGCTGACCGCAAGACGCTCGTCGCCA 120
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 140 GCCTGCATGTTCGGCGAGCCCATCAGCTGCTCGACGCTGACCGCAAGACGCTCGTCGCCA 199
OY 121 CCATCGAGTACCTGGTTCGCTCGACGACGCTCTCAGGCTGGCCAGGCCCGCTTATGA 180
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 200 CCATCGAATATCTGCTCCGCTTCACGAGG-----GTCAGACCACGATGA 244
OY 181 CTGTCCCGCGCGGTCGAGGTGCGCGTGGAAACCGACGACAT 223
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 245 CCGTTCCGCGCGGTCGAGGTGCGCGTGGAAACCGACGACAT 287
```

```

/transcript_cab1
/product="RNA polymerase beta subunit"
/protein_id="CAC87032.1"
/db_xref="GI:22208407"

```

translation="TDEALLDIYVKLRGCEPTKESAOITLLENLFKEKRYRDLANVGR  
YKVNKKLIVGEPTISSTLTEDVATIEYLVRLHGCEQTMTPGVGEVETDIDD  
HFGNKKRLTIVGELIQNIQVGSRMKRVYRMTTQDVEAITPKTLINIRPVAAIKI  
FFGTQSQPMQNNPLSGITLTKRKLIALGPGGLSRERAGLEVYRD"

55.5%; Score 123.8; DB 1; Length 618;  
 tity 76.7%; Pred. No. 8.7e-14;  
 conservative 0; Mismatches 37; Indels 15; Gaps 1;

CGGCTACGACCTGGCCCGCTCGGCCGCTACAAGGTCAACAAGAGTGG 60

CGATCATTCGATACCCACCAGCAGCTGACCGAAGAAGACCTGTGCCECA 120  
|||||  
GGCGGAGCCCATCACGTCGTGAGCGCTGACCGGAAGAAGACGCTGTGCCECA 214

CTGGTTCGGCTGCACACGCGCTCTCAGGGTGCCAGCCCCCGTTATGA 180  
||||| ||| ||||| | | |||  
CTGGTTCGGCTGCACGAGG- - - - -GTCAGACACCGATGA 259

GGGGTCGAGGTGCGGTGGAAACCGACACAT 223  
 GGGGTCGAGGTGCGGTGGAAACCGACACAT 302

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

7 618 bp DNA circular BCT 09-AUG-2002  
Mycobacterium tuberculosis partial rpoB gene for RNA polymerase  
unit, isolate 2348-98.

1 GI:22208410  
 erase beta subunit; rpoB gene.  
 rium tuberculosis.  
 rium tuberculosis.

Actinomyces, tuberculousis  
Actinobacteria; Actinobacteridae; Actinomycetales;  
Actinomycetaceae; Actinomycetidae; Actinomycetinae;  
Mycobacteriaceae; Mycobacterium; Mycobacterium  
tuberculosis complex.

..., Jimenez, M.S. and Saez, J.A. Analysis of rifampin-resistant *Mycobacterium tuberculosis* in Spain (1996-2001). Description of new alleles into *rpoB*.

review  
ed  
s 1 to 618)

id (07-AUG-2001) Herrera L., Servicio Bacteriología, Centro Microbiología., Ctra. Majadahonda-Pozuelo, Km 2.5, Majadahonda. Madrid. 28220. SPAIN

```
Location/Qualifiers
1. 618
/organism="Mycobacterium tuberculosis"
/isolate="2348-98"
```

```
/db_xref="taxon:1773"  
1. .618  
/gene="rpoB"  
<1. .618
```

```
/gene="rpoB"  
/codon_start=1  
/transl_table=11  
/product="RNA polymerase beta subunit"
```

```
/protein_id="CAC87034.1"  
/db_xref="GI:22208411"
```

```
1. .615
/ gene="rpoB"
<1. .>615
```

```

/gene="rpoB"
/codon_start=1
/transl_table=11
/product="RNA polymerase beta subunit"
/protein_id="CAC87030.1"

```

/db\_xref- "GI:222084303"  
 /translation- "LDIVRKLPPEPTKSAOTLLENLFFKEKRYDLARGVRYKYNK  
 KRLGVGEIPTSSTLEEDVATIEYLVRHCEGTTTPGVGEVPEVTDIDHFCNR  
 KLHTVGEILONOIVGMSMRVYKEMHTQDVEAITPTOLNIRPVAAKEFFGTS  
 QLSQPIFMQNNPLSLGTHIKRRLSALGPGGLSREKAGLEVLRDHP"

191 C 201 q 99 L  
55.5%; Score 123.8; DB 1; Length 616;  
76.7%

Identity 76.7% pred. NO. 8/e-14;  
 conservative 0; Mismatches 37; Indels 15; Gaps 1;  
 TGGCTACGACCTGGCCCGCTCGCCGCTACACGCTCACACAGAGCTGG 60

CGCTACGACGTGGCCCGCTCGTATAGGTCAACAAGAGTCG 139  
GATCATCCGA"CAACCACCGACGC"TGACCCGAGAAAGACGTCGTCGCA 120

GGGAGGCCATCACCTCTGCGCGCTGACCGAAGAGACGTCGTGGCCA 199  
CTGTTTCGCTGGACACAGCCCTCTCAGGTGGCCAGGCCCGCTTATGA 180

CTGGTCCGCTTGCACGAG- - - - - GTCAGACCACCGATGA 244  
GGGGTTCGAGGTGCCGGTGTGAAACCGACGACAT 223

GGCGTCGAGGTGCCGTGGAAACCGACACAT 287

618 bp DNA circular BCT 09-AUG-2002  
Mycobacterium tuberculosis partial rpoB gene for RNA polymerase

1 GI:22208406  
merase beta subunit; rpoB gene.  
unit, isolate 2540-97.

Actinobacteria; Actinobacteridae; Actinomycetales; Actinomycetaceae; Mycobacteriaceae; Mycobacterium tuberculosis; Mycobacterium tuberculosis

... Jimenez, M. S. and Saez, J. A.  
... analysis of rifampin-resistant *Mycobacterium tuberculosis*  
... complex.

in Spain (1996-2001). Description of new alleles into rpoB review

Submission  
(07-AUG-2001) Herrera L., Servicio Bacteriología, Centro  
Microbiología, Ctra. Matadabonda-Dozuelo, km. 25

da. Madrid. 28220, SPAIN  
Location/Qualifiers  
1. 618  
/organism="Mycobacterium tuberculosis"

```

/0-5a10m" type="text" value="1773"
/isolate="2540-97"
/db_xref="taxon:1773"
1. .618
/geom="rect"

```

```

/genes-1.pdb
<1. .>618

```

/translation="TDEALLDIYKLRGPPTKESAOQLLENLFFKRYDLARVGR  
YKVNKGLHVGEDITSTLDEEDVATIEYVRLHHEGQTTMTVPGGVEVPVETDDID  
HFGNRRURTVELGIONQIRVMSRMERVVRMTTQDVEAITPTQLINIRPVVAAIKE  
FFGTSQSQSQFGMGNPLSLGTHKRLSALPGGLSRERAGLEVRDV"  
HASE COUNT 124 a 192 c 207 g 95 t  
ORIGIN

Query Match 55.5%; Score 123.8; DB 1; Length 618;  
Best Local Similarity 76.7%; Pred. No. 8.7e-14;  
Matches 171; Conservative 0; Mismatches 37; Indels 15; Gaps 1;  
  
QY 1 TCAAGGAGAGCGCTACGACCTGGCCCGCGCTGCGGCTACAAGGTCAACAAGAGCTGG 60  
DB 95 TCAAGGAGAGCGCTACGACCTGGCCCGCGCTGCGGCTACAAGGTCAACAAGAGCTGG 154  
QY 61 GCCTGAACACCGATCATCCGATCACCAACGACGCTGACCGAAGAACGCTGTCGCCCA 120  
DB 155 GCGTCGATGTCGGCGAGCCCATCAGCTGTCGACGCTGACCGAAGAACGCTGTCGCCCA 214  
QY 121 CCATCGAGTACTGCTGCTGCGACCAACGCTCTCAGGTTGGCCAGGCCCGCTTATGA 180  
DB 215 CCATCGAATATCTGGTCCGCTTGCACGAGG-----GTACAGCACCGATGA 259  
QY 181 CTGTCCCGGGGGTCTGAGGTGCGGCTGGAACCGACGACAT 223  
DB 260 CCGTTCGGGGCGGCTGAGGTGCGGCTGGAACCGACGACAT 302

RESULT 6  
MTU318814  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
  
MTU318814  
Mycobacterium tuberculosis partial rpoB gene for RNA polymerase  
beta subunit, isolate 1058-97.  
AJ318814  
AJ318814.1 GI:22208404  
RNA polymerase beta subunit; rpoB gene.  
Mycobacterium tuberculosis.  
Mycobacterium tuberculosis.  
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium  
tuberculosis complex.

REFERENCE  
AUTHORS  
TITLE  
  
Herrera, L., Jimenez, M.S. and Saez, J.A.  
Molecular analysis of rifampin-resistant Mycobacterium tuberculosis  
isolated in Spain (1996-2001). Description of new alleles into rpoB  
gene and review  
Unpublished  
2 (bases 1 to 633)  
Herrera, L.  
Direct Submission  
Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro  
Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5,  
Majadahonda, Madrid. 28220, SPAIN

FEATURES  
source  
1..633  
/organism="Mycobacterium tuberculosis"  
/isolate="1058-97"  
/db\_xref="taxon:1773"  
1..633  
/gene="rpoB"  
/gene="rpoB"  
<1..>633  
/codon\_start=1  
/transl\_table=11  
/product="RNA polymerase beta subunit"  
/protein\_id="CAC87031.1"  
/db\_xref="GI:22208405"  
/translation="DNTVGTDEALLDIYKLRPGSPPTKESAOQLLENLFFKRYDL  
ARVGRYKVNKGLHVGEDITSTLDEEDVATIEYVRLHHEGQTTMTVPGGVEVPVE  
TDDIDHFGNRRURTVELGIONQIRVMSRMERVVRMTTQDVEAITPTQLINIRPVV  
AAIKEFFGTSQSQSQFGMGNPLSLGTHKRLSALPGGLSRERAGLEVRDV"  
BASE COUNT 129 a 195 c 210 g 99 t  
ORIGIN

Query Match 55.5%; Score 123.8; DB 1; Length 633;  
Best Local Similarity 76.7%; Pred. No. 8.7e-14;  
Matches 171; Conservative 0; Mismatches 37; Indels 15; Gaps 1;  
  
QY 1 TCAAGGAGAGCGCTACGACCTGGCCCGCGCTGCGGCTACAAGGTCAACAAGAGCTGG 60  
DB 110 TCAAGGAGAGCGCTACGACCTGGCCCGCGCTGCGGCTACAAGGTCAACAAGAGCTGG 169  
QY 61 GCCTGAACACCGATCATCCGATCACCAACGACGCTGACCGAAGAACGCTGTCGCCCA 120  
DB 170 GCGTCGATGTCGGCGAGCCCATCAGCTGTCGACGCTGACCGAAGAACGCTGTCGCCCA 229  
QY 121 CCATCGAGTACTGCTGCTGCGACCAACGCTCTCAGGTTGGCCAGGCCCGCTTATGA 180  
DB 230 CCATCGAATATCTGGTCCGCTTGCACGAGG-----GTACAGCACCGATGA 274  
QY 181 CTGTCCCGGGGGTCTGAGGTGCGGCTGGAACCGACGACAT 223  
DB 275 CCGTTCGGGGCGGCTGAGGTGCGGCTGGAACCGACGACAT 317

RESULT 7  
MTU318816  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
  
MTU318816  
Mycobacterium tuberculosis partial rpoB gene for RNA polymerase  
beta subunit, isolate 1255-98.  
AJ318816  
AJ318816.1 GI:22208408  
RNA polymerase beta subunit; rpoB gene.  
Mycobacterium tuberculosis.  
Mycobacterium tuberculosis.  
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium  
tuberculosis complex.

REFERENCE  
AUTHORS  
TITLE  
  
Herrera, L., Jimenez, M.S. and Saez, J.A.  
Molecular analysis of rifampin-resistant Mycobacterium tuberculosis  
isolated in Spain (1996-2001). Description of new alleles into rpoB  
gene and review  
Unpublished  
2 (bases 1 to 637)  
Herrera, L.  
Direct Submission  
Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro  
Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5,  
Majadahonda, Madrid. 28220, SPAIN

FEATURES  
Location/Qualifiers  
1..637  
/organism="Mycobacterium tuberculosis"  
/isolate="1255-98"  
/db\_xref="taxon:1773"  
1..637  
/gene="rpoB"  
/gene="rpoB"  
<1..>637  
/codon\_start=1  
/transl\_table=11  
/product="RNA polymerase beta subunit"  
/protein\_id="CAC87033.1"  
/db\_xref="GI:22208409"  
/translation="VGTDEALLDIYKLRPGSPPTKESAOQLLENLFFKRYDLARV  
GRYKVNKGLHVGEDITSTLDEEDVATIEYVRLHHEGQTTMTVPGGVEVPVETDD  
IDHFGNRRURTVELGIONQIRVMSRMERVVRMTTQDVEAITPTQLINIRPVVAAI  
KEFFGTSQSQSQFGMGNPLSLGTHKRRFALPGGLSRERAGLEVRDVHPSH"  
BASE COUNT 128 a 198 c 210 g 101 t  
ORIGIN

Query Match 55.5%; Score 123.8; DB 1; Length 637;  
Best Local Similarity 76.7%; Pred. No. 8.7e-14;  
Matches 171; Conservative 0; Mismatches 37; Indels 15; Gaps 1;  
  
QY 1 TCAAGGAGAGCGCTACGACCTGGCCCGCGCTGCGGCTACAAGGTCAACAAGAGCTGG 60



```

101 TCAAGGAGAGCGCTAGGACCTGGCCCGCTGCTCGCTATAGGTACACAAGAGCTCG 160
|||||
61 GCCTGAACACCGGATCATCGGATCACCACGACGCTGACCGAAGAGACGCTGTCGCCCA 120
|||||
161 GCGTGCATGTCGCGGAGCGCATCACGTCGTCGACGCTGACCGAAGAAGACGCTGTCGCCCA 220
|||||
121 CATCGAGTACCTGGTTCCCTGTCACACGCGCTCTCAGGTGGCCAGGCCCGCTTATGA 180
|||||
221 CCATCGAATATCTGTGCTCGCTTCACGAGG-----CTCAGACACGACATGA 265

181 CTCTCCCGCGGGGTCGAGGTGCGGTCGGAACCGACGACAT 223
|||||
266 CCTCTCCGCGCGCTCGAGGTGCGGTCGGAACCGACGACAT 308

RESULT 8
LOCUS MTU318821 639 bp DNA circular BCT 09-AUG-2002
DEFINITION Mycobacterium tuberculosis partial rpoB gene for RNA polymerase
beta subunit, isolate 1071-98.
ACCESSION AJ318821
VERSION AJ318821.1 GI:22208418
KEYWORDS RNA polymerase beta subunit; rpoB gene.
SOURCE Mycobacterium tuberculosis.
ORGANISM Bacteria; Actinobacteria; Actinomycetales;
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
tuberculosis complex.
1 Herrera,L., Jimenez,M.S. and Saez,J.A.
REFERENCE Molecular analysis of rifampin-resistant Mycobacterium tuberculosis
AUTHORS isolated in Spain (1996-2001). Description of new alleles into rpoB
TITLE gene and review
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 639)
AUTHORS Herrera,L.
JOURNAL Direct Submission
TITLE Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro
National Microbiologia, Ctra. Majadahonda-Pozuelo, Km 2.5,
Majadahonda. Madrid. 28220, SPAIN
FEATURES Location/Qualifiers
source
1..639
/organism="Mycobacterium tuberculosis"
/isolate="1071-98"
/db_xref="taxon:1773"
1..639
/gene="rpoB"
<1..>639
/gene="rpoB"
/codon_start=1
/transl_table=11
/product="RNA polymerase beta subunit"
/protein_id="CAC87038.1"
/db_xref="GI:22208418"
/translation="RRTDEALLDIYKRLPGEPPTKESAOTLLENLFFKEKRYDLARV
GRYKNKGLGIHVGEPIFSTLTEDVVATIEYLVRHGEQTTWVPGVEVPVETDD
IDHGNRLRTVVGELIONIRVGMKRVYREKMTQDVDAITPQTLINIRPAVAAI
KEFTCSLSQFMVQNPPLSLGTRKRRLSALGPGLSRERAGLEVDRVHPSHS"
BASE COUNT 126 a 202 c 212 g 99 t
ORIGIN
Query Match 55.5%; Score 123.8; DB 1; Length 639;
Best Local Similarity 76.7%; Pred. No. 8.7e-14;
Matches 171; Conservative 0; Mismatches 37; Indels 15; Gaps 1;

QY 1 TCAAGGAGAGCGCTAGGACCTGGCCCGCTGACAAAGTCAACAAGAGCTGG 60
|||||
DB 101 TCAAGGAGAGCGCTAGGACCTGGCCCGCTGCTATAGGTACACAAGAGCTCG 160
|||||
QY 61 GCCTGAACACCGGATCATCGGATCACCACGACGCTGACCGAAGAGACGCTGTCGCCCA 120
|||||
DB 101 TCAAGGAGAGCGCTAGGACCTGGCCCGCTGCTATAGGTACACAAGAGCTCG 160
|||||
QY 61 GCCTGAACACCGGATCATCGGATCACCACGACGCTGACCGAAGAGACGCTGTCGCCCA 120
|||||
DB 161 GCGTGCATGTCGCGGAGCGCATCACGTCGTCGACGCTGACCGAAGAAGACGCTGTCGCCCA 220
|||||

```

```

QY 121 CCATCGAGTACCTGGTTCCCTGTCACACGCGCTCTCAGGGTGGCCAGGCCCGCTTATGA 180
|||||
DB 221 CCATCGAATATCTGTGCTCGCTTCGACGAGG-----GTCAGACACGACATGA 265
|||||
QY 181 CTGTCCCGCGGGTTCGAGGTGCGGTCGGAACCGACGACAT 223
|||||
DB 266 CCGTTCGCGCGCGCTCGAGGTGCGGTCGGAACCGACGACAT 308
|||||

RESULT 9
LOCUS 150706 970 bp DNA linear PAT 07-UCT-1997
DEFINITION Sequence 1 from patent US 5643723.
ACCESSION 150706
VERSION 150706.1 GI:2472409
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 970)
AUTHORS Persing,D.H., Hunt,J.J., Young,K.K.Y., Felmler,T.A., Roberts,G.D.
and Whelan,A.Christian.
TITLE Detection of a genetic locus encoding resistance to rifampin in
mycobacterial cultures and in clinical specimens
JOURNAL Patent: US 5643723-A 1 01-JUL-1997;
FEATURES Location/Qualifiers
source
1..970
/organism="unknown"
BASE COUNT 182 a 302 c 330 g 156 t
ORIGIN
Query Match 55.5%; Score 123.8; DB 6; Length 970;
Best Local Similarity 76.7%; Pred. No. 8.3e-14;
Matches 171; Conservative 0; Mismatches 37; Indels 15; Gaps 1;

QY 1 TCAAGGAGAGCGCTAGGACCTGGCCCGCTGCGCGCTACAAAGTCAACAAGAGCTGG 60
|||||
DB 26 TCAAGGAGAGCGCTAGGACCTGGCCCGCTGCTATAGGTACACAAGAGCTCG 85
|||||
QY 61 GCCTGAACACCGGATCATCGGATCACCACGACGCTGACCGAAGAGACGCTGTCGCCCA 120
|||||
DB 86 GCGTGCATGTCGCGGAGCGCATCACGCTGTCGACGCTGACCGAAGAGACGCTGTCGCCCA 145
|||||
QY 121 CCATCGAGTACCTGGTTCCCTGTCACACGCGCTCTCAGGGTGGCCAGGCCCGCTTATGA 180
|||||
DB 146 CCATCGAATATCTGTGCTCGCTTGCACGAGG-----GTCAGACACGACATGA 190
|||||
QY 181 CTGTCCCGCGGGTTCGAGGTGCGGTCGGAACCGACGACAT 223
|||||
DB 191 CCGTTCGCGCGCGCTCGAGGTGCGGTCGGAACCGACGACAT 233
|||||

RESULT 10
LOCUS AX111339 3534 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 2072 from Patent W00123604.
ACCESSION AX111339
VERSION AX111339.1 GI:13927631
KEYWORDS
SOURCE Mycobacterium tuberculosis.
ORGANISM Mycobacterium tuberculosis
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Corynebacterineae; Mycobacteriaceae;
Mycobacterium; Mycobacterium tuberculosis complex.
REFERENCE 1 (bases 1 to 3534)
AUTHORS Bergeron,M.G., Boissinot,M., Huletsky,A., m Nard,C., Ouellette,M.,
Picard,F.J. and Roy,P.H.
TITLE Highly conserved genes and their use to generate probes and primers
for detection of microorganisms
JOURNAL Patent: W0 0123604-A 2072 05-APR-2001;
FEATURES Infectio Diagnostic (I.D.I.) INC. (CA)
Location/Qualifiers

```

source	1..3534
BASE COUNT	679 a 1081 c 1188 g 586 t
ORIGIN	<p>MSRERVVRVMTTQDVEAITPOTLINIRPVAAIKIEFFCTQSLSQFMDONNPISGLG  HKRRLSALPGGLSLRERAGLEVDRVHPSHYGRMCPIETPEGPNIIGLISLSVYARVKN  FGFETPYRKVVDGWSDGIVYLTADEEDRHVVQAQNSPIDADGRVEPRVLVRKKAQ  EVEYVPSSEVDYMDVSPROMYSVATAM1PFEHDDANRALMGANMOROQAVPLVRSEAP  LVGTGMELRAALDAATSSQESGVIEVSADYITVHDNGTRRTYIKRKFARNSHUGTC  ANQCP1VDAGDRVEAGQVIADGCTDDGEMALGNLLUVAIMPWEHNGEDATILSNRL  VEEDVLSHIEHEHEDAROTKLGAEIITRDIPNTIDEVLADLDERGIVIRI GAEDVRS  DILYGVTKPGTELTPEERLLRAIFEGEAREVDRITSLKVPHGSGKVICIRVFSRED  EDELPAQVNELVRYVAQKRKISDGLAGRHGKVGIGKILPVEDOMFLADCTPVDI  ILNTHGVPRRMNIGQIILETHLGMCAHSGKVDAAKGVPMWAARLPDELLEAOPNAIVS  TPVFDGAQAEALQELLSCITLPNRDGDVLDADGKAMLFDCRSGEPFPVTVGYMYIM  KLHLVDDKIHARSTGPTSMITQOPLCGKAQFGQGRFEGMECWAMQAYCAATYIQLLEL  TIKS"</p>
Query Match	55.5%; Score 123.8; DB 6; Length 3534;
Best Local Similarity	76.7%; Pred. No. 7.2e-14;
Matches 171; Conservative	0; Mismatches 37; Indels 15; Gaps 1;
BASE COUNT	723 a 1173 c 1293 g 664 t
ORIGIN	<p>1 TCAAGGAGAACGCTACGACCTGCGCCGCTCGCGCGCTACAAGTCAACAGAAGCTGG 60   Db 902 TCAAGGAGAACGCTACGACCTGCGCCGCTCGGTGCGTATTAAGTCAACAGAAGCTCG 961   Qy 61 GCCTGAACACCGATATCCGGATCACCAACACGACGCTGACCGAAGAAGCGTCGTCGCCA 120   Db 962 GGCTGCATGTCGCGAGCCCATACGCTCTCGACGCTGACCGAAGAAGCGTCGTCGCCA 1021   Qy 121 CCATCGAGTACCTGGTTCGGCTGCACCAACCGCCCTCTCAGAGTGGCCAGGCCCGCGTTATGA 180   Db 1022 CCATCGAATATCTGGTTCGGCTGTCACGAGG-----GTCAGACCCACGATGA 1066   Qy 181 CTGTCCCGCGCGGCTCGAGGTCGCGGTGGAACACGACGACAT 223   Db 1067 CCGTTCGGCGCGGCTCGAGGTCGCGGTGGAACACGACGACAT 1109   </p>
RESULT 11	MTU12205
LOCUS	3853 bp DNA linear BCT 02-MAR-2000
DEFINITION	Mycobacterium tuberculosis H37Rv RNA-polymerase beta subunit (rpoB) gene, partial cds.
ACCESSION	U12205
VERSION	U12205.1
KEYWORDS	GI:515684
SOURCE	Mycobacterium tuberculosis.
ORGANISM	Mycobacterium tuberculosis Bacteria; Firmicutes; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium tuberculosis complex.
REFERENCE	1 (bases 1 to 3853) Imboden, P., Troller, R., Marchesi, F., Telenti, A., Bodmer, T., Cole, S., Schopfer, K. and Burkart, T. The rpoB gene of Mycobacterium tuberculosis Unpublished
AUTHORS	Imboden, P., Troller, R., Marchesi, F., Telenti, A., Bodmer, T., Cole, S., Schopfer, K. and Burkart, T.
TITLE	The rpoB gene of Mycobacterium tuberculosis
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 3853) Imboden, P.
AUTHORS	Imboden, P.
TITLE	Direct Submission
JOURNAL	Submitted (11-JUL-1994) Paul Imboden, Institute for Medical Microbiology, University of Berne, Friedbuehlstrasse 51, Berne, 3010, Switzerland
FEATURES	Location/Qualifiers
source	1..3853 /organism="Mycobacterium tuberculosis" /strain="H37Rv" /db_xref="taxon:1773" 576..>3853 /gene="rpoB" 576..>3853 /gene="rpoB" /codon_start=1 /transl_table=11 /product="RNA-polymerase beta subunit" /protein_id="AAA20242.2" /db_xref="GI:7144499" /translation="MLEGCLADRSQKTAASPSRQSSNNNSVPGAPNRVSPFAKL RLEPVGLLDVQDSFELWLGSPRWREAAERGDVNPVGLVEYLSPIEDFSGS NSLSDFRFDVKAQVDECKDQMTAAFLVTAEP INNNTEIKSPVFMGDFPM TEKGTFTIIRVSVOLVSPGVYFDETDKSTKTLHSKVIIPSRGAMFEFDVK RDYGVIRIDRRKRPVTVLLKALGKTSQIVVERGFSEIMRSTLEKDNVTGDEALLD LYRKLRCPEPTKSAQOTLLENLFFKCKRDLARVGRVKNKGLHVGEPITSTLT EDVWVATIEVLRHVGQTMVPGGVEVPEVDDIDHFNRRRLRTVGLIQNQRVIG
gene	1..3853 /organism="Mycobacterium tuberculosis" /strain="H37Rv" /db_xref="taxon:1773" 576..>3853 /gene="rpoB" 576..>3853 /gene="rpoB" /codon_start=1 /transl_table=11 /product="RNA-polymerase beta subunit" /protein_id="AAA20242.2" /db_xref="GI:7144499"
CDS	1..3853 /organism="Mycobacterium tuberculosis" /strain="H37Rv" /db_xref="taxon:1773" 1065..4598 /gene="rpoB" 1065..4598 /gene="rpoB" /codon_start=1 /evidence=experimental /transl_table=11 /product="RNA-polymerase beta subunit" /protein_id="AAA21416.1" /db_xref="GI:468334" /translation="MLEGCLADRSQKTAASPSRQSSNNNSVPGAPNRVSPFAKL



```
SPASGRPLAMPRLUMVTG1XYLTTEVPEDTCEYQPASGDHPETCVYSSPAFAIMAADR
GVLSYRAK1KVRULTOLRPVETIAELFGHSGWQPDAGWMAETTLGRVMFNELLPLGTP
FVNQKHKKVOAA1IINDLAERYPMIVVAQTVDKLQDAGFVWATRSRGTVSMADLVPPP
RKKEILDHYEADKVEKQFQKGLNHNDEALVEITWKEATDEVGQALREHYPPDNP
1ITITDVSQATCNFTQFTLAGMGLVTNPKGEFIPRPVKSFRGLTVLVEYINTHCA
RKGLADTALRTADSGYLTRLIVUSODVJVRHDCOTERGIVVELAERAPDGLTIRDP
YLTSAVARTLCTDPADEACNVIVRGODLGDPEIDALLAAGTIVQVRSVLYTCATST
GVATCVRSMATCKLVDIGEAVGIVAAQSIGEPGTOLTMRFTHQGVGVEDITGGGLPR
VLELFEARVPRGRAPTADTVTQGRVRLDEGERFYKITIVPDGGEVYVDKISKORLVR
FKHEDGSERVLSGDHVEVGGQMEGSDHEVLRVQGPPEVQIHLVREVOEYVRAQG
VSIHDKIEVIVQMLRRVTIIDSQSTFELPGSLIDRAEFAENRRVAVAGGPPAAGR
PVLWGLITFKASLATSLSAASFOFTTHVLDDAALNCRSDKLNGLKENVLIIGKLI PACT
G1NRYNIAVQPTFEARAAAYTIPSYEDQIYSPDFGAGTAAVPLDDYGYSDYR"
complement(7691..8065)
/gene="MT0697"
complement(7691..8065)
/gene="MT0697"
/notes="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAK44923.1"
/db_xref="GI:13880220"
/transl_table="MT0697"
/translation="MFDSSAAITNPGHAWASAMERSGLLECVACLDDEQPFGETADKL
NPDKGSSRRVPRQADCGIATHVERGGQSGQGVVQRMHGGPFPALAMQDRLIHH
GEQTONKIAQAFRRFCVCSPT"
complement(8058..9972)
/gene="MT0698"
/notes="This region contains an authentic frame shift and
is not the result of a sequencing artifact; identified by
Glimmer2; putative; conserved hypothetical protein,
authentic frameshift"
10167..10925
/gene="MT0699"
10167..10925
/gene="MT0699"
/notes="identified by match to PfAM protein family HMM
PF01261"
/codon_start=1
/transl_table=11
/product="AP endonuclease, family 2"
/protein_id="AAK44924.1"
/db_xref="GI:13880221"
/transl_table="MLGSHVSPTUPLAAAEAGADVWQIFLGNQSNKAPKPRDDAA
ALKAATLPTVTHAPYLINLASANNRVRIPSRKILQETCAAAADIGAAAVIVHGGHVA
DNIDKGFQWRKALDRLETEVPVYLENTAGGDHMAARREDTIARLMDVIGDTGIGFC
LDTCHTWAAGEALTDAVDRIKAITGRIDLVCNDSDEAGSGRDRHANLGSQIDPDL
LVAAVKAAGAPVICETADQGRKDDIAPLERTGS"
10957..11799
/gene="MT0700"
10957..11799
/gene="MT0700"
/notes="similar to GB:U00012 PID:466863; identified by
sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="hydrolase/esterase, putative"
/protein_id="AAK44925.1"
/db_xref="GI:13880222"
/transl_table="MLRRVA1LILAAVLAFAGCSGCTRLAAGFGNGSVHPTLDVDGAGR
SYRLYKPVGLPSSAPLVMLHGGFGSAKOAESYGMDELADSEKFLVAYPDGYHRAWN
ANGGCCGRPARBEGVDVIGYVRAVADIANVSDIPARVYVYTGSHNGAINSYTLACNT
SIFPAIGVSGTQDPCQSPRPVSVIHGTADPLVRYHGGPGAGARIDGPPVPDLN
AFWRNRCGALDPTTEGTPVTSGATCADNRNRRVLLTVDDAGHRWPSFATQTLWRFFA
AHR"
11859..13487
/gene="MT0701"
11859..13487
/gene="MT0701"
/notes="similar to SP:P33224 GB:L20915 PID:457172
PID:457174 PID:537028; identified by sequence similarity;
putative"
/codon_start=1
```

```
/transl_table=11
/product="acyl-CoA dehydrogenase, putative"
/protein_id="AAK44926.1"
/db_xref="GI:13880223"
/transl_table="MSDTHVVTNVPPLNYPASSPVLIETALIOEGCGWGLDEVNEV
GALSASQORMCCELADNRNPILHTHDAYGYRVEVEYDPAYHELAKRTAITHCMHAAP
WADRLPATTAKGITACITAMSTERGSDVRAGTTOATPNADGYSLTGRKWFTSAPMCD
DPELKPATTKAGTACITAMSTERGSDVRAGTTOATPNADGYSLTGRKWFTSAPMCD
IFLVLQAQPDGLSCFLPRVLPDGTNRNMFQRLKDKLGNHANASSEVEYDCAVAWLV
GEGRCVPTTIEMVNLTRLDCAIGSATSMTGLTRAVHHAQHRKAFCAVLIIDQPLMRN
VLADLAEAEAAITVAMRAGATDNVRCNETEALLRRLGLAAAKYVWCKRSFAHAAE
ALPCLGNGYVEDSGMPRLYREAPLNGIMEGSGNVSLDITLRAMATRPACVVELFDEL
ARSAGODPLDGHVERLRQLDGLDTIGYRARKIAEDICLALOGSLLVRHGHFAVAEA
FIATRLGGWGWGAYGTMPTAGLDIAPILERALVAG"
13498..14436
/gene="MT0702"
13498..14436
/gene="MT0702"
/notes="similar to GP:3885480; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="enoyl-CoA hydratase/isomerase family protein"
/protein_id="AAK44927.1"
/db_xref="GI:13880224"
/transl_table="MTHAIRPVDFDNLKMTYEVVTGRIARITFNRPEKGNALIIDTPLL
ELSAVERADLDGPHVHILVSGREGFCAGFDLSAYAGSSSTGGGAYOGTVLDGKT
QVANHLPNPNQWDPMDIYOMMSRFRVGRFASLHMADKPTVVKIHGCVAGGOTDIALHADQ
YVAAADAKIGYDPTTRVGMVPAAGLWHRIGLDQRAKRLFTGDCITCAOAAEAGLAVEA
PEPADLDERTERLVARIALPVNQLIMVKLALNSALLQGVATSRMVSTVFDCAARHT
PEGHAFVADAVEHGFPRDAVRRRDEPDGYGQASRV"
14439..15161
/gene="MT0703"
14439..15161
/gene="MT0703"
/notes="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
Query Match 55.5%; Score 123.8; DB 1; Length 19352;
Best Local Similarity 76.7%; Pred. No. 6e-14;
Matches 171; Conservative 0; Mismatches 37; Indels 15; Gaps 1;
QY 1 TCAAGGAGAACGGCTACGACCTGGCCGCGCTCGGCCGCTACAAGGTCAACAAGAGCTGG 60
Db 1064 TCAAGGAGAACGGCTACGACCTGGCCGCGCTCGGCCGCTACAAGGTCAACAAGAGCTGG 1123
QY 61 GCCTGAACACCGGATCATCCGATCACCACGACGCGCTGACCGGAAGAGAGCTGTCGCCCA 120
Db 1124 GGCTGCATGTGCGGCGAGCCCATCAGCTCGTGACGCTGACCGAAGAGAGCTGTCGCCCA 1183
QY 121 CCATCGAGTACCTGGTTCGCGCTGCACCACGCGCTCTCAGGCTGGCCAGCCCGCTTATGA 180
Db 1184 CCATCGAATATCTGGTCCGCTTCACAGG-----GTCACACCACCATGA 1228
QY 181 CTGTCGCCGCGGGTCGAGGTGCCGCTGGAACCCGACGACAT 223
Db 1229 CCGTTCGCGGCGGCTCGAGGTGCCGCTGCAACCCGACGACAT 1271
RESULT 14
MTC1376
LOCUS
DEFINITION
Mycobacterium tuberculosis H37Rv complete genome; segment 32/162.
ACCESSION
295972 AL123456
VERSION
295972.1 GI:3261790
KEYWORDS
Mycobacterium tuberculosis H37Rv.
Mycobacterium tuberculosis H37Rv.
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Corynebacterineae; Mycobacteriaceae;
Mycobacterium tuberculosis complex.
```

REFERENCE  
AUTHORS

1 (bases 1 to 19770)  
Cole,S.T., Brosch,R., Parkhill,J., Garnier,T., Churcher,C.,  
Harris,D., Gordon,S.V., Eigmeier,K., Gas,S., Barry III,C.E.,  
Tekala,F., Badcock,K., Basham,D., Brown,D., Chillingworth,T.,  
Connor,R., Davies,K., Devlin,K., Feltwell,T., Gentles,S.,  
Hamlin,N., Holroyd,S., Hornsby,T., Jagels,K., Krogh,A., McLean,J.,  
Moule,S., Murphy,L., Oliver,S., Osborne,J., Quail,M.A.,  
Rajandream,M.A., Rogers,J., Rutter,S., Seeger,K., Skelton,S.,  
Squares,S., Squires,R., Sulston,J.E., Taylor,K., Whitehead,S. and  
Barrell,B.G.  
Deciphering the biology of Mycobacterium tuberculosis from the  
complete genome sequence  
Nature 393 (6685), 537-544 (1998)  
98295987  
9634230  
2 (bases 1 to 19770)  
Parkhill,J.  
Direct Submission  
Submitted (11-JUN-1998) Submitted on behalf of the Mycobacterium  
tuberculosis sequencing and mapping teams, Sanger Centre, Wellcome  
Trust Genome Campus, Hinxton, Cambridge CB10 1SA Unite de Genetique  
Moleculaire Bacterienne, Institut Pasteur, 28 rue du Docteur Roux,  
75724 Paris Cedex 15, France E-mail: parkhill@sanger.ac.uk  
On Jun 27, 1998 this sequence version replaced gi:2143285.

## COMMENT

Details of M. tuberculosis sequencing at the Sanger Centre are  
available on the World Wide Web.  
(URL, [http://www.sanger.ac.uk/Projects/M\\_tuberculosis/](http://www.sanger.ac.uk/Projects/M_tuberculosis/)) CDS have  
been renumbered from the original cosmid submissions but the old  
gene designations are in brackets after the new gene numbers.  
Gene prediction was based on a Hidden Markov Model of TB genes  
implemented in Ttparse (Krogh) supplemented with visual inspection  
of positional base preference in codons, especially where there is  
an increase in the observed/expected third position G + C.  
CAUTION: In some cases we may not have predicted the correct  
initiation codon. Where possible we choose an initiation codon  
(atg, gtg, or ttg) which is preceded by an upstream ribosome  
binding site sequence (optimally 5-13bp before the initiation  
codon). If this cannot be identified we choose the most upstream  
initiation codon.

FEATURES  
Location/Qualifiers

1..19770  
/organism="Mycobacterium tuberculosis H37Rv"  
/strain="H37Rv"  
/db\_xref="taxon:83332"  
/clone="I376"  
57..61  
/note="agga, possible rbs upstream of Rv0654"  
68..1573  
/gene="Rv0654"  
68..1573  
/gene="Rv0654"  
/note="Rv0654, (MTC1376.22), len: 501. unknown, FASTA  
score: Q5353 LIGNOSTILBENE-ALPHA,BETA-DIOXYGENASE (485  
aa) opt:280 z-score: 330.1 E(): 2.3e-11, (28.5% identity  
in 523 aaoverlap). Also similar to M. tuberculosis protein  
MTCY21C12.07c (29.5% identity in 522 aa overlap)"  
/codon\_start=1  
/transl\_table=11  
/product="hypothetical protein Rv0654"  
/protein\_id="CAB09380.1"  
/db\_xref="GI:2143307"  
/db\_xref="SPTREMBL:O06785"

## RBS

## gene

## CDS

/translation="MTTAAAESNPYLEGLFAPVSTEVATDLPVTCRIPEDHLDGRY  
LRNGPNPAEVDPATYHWFTGDAMVHGVALRDGRARWTRNRVTRPACVAAALGEPISA  
RHPQRTGIIEGGPNTNLTAGRTALVAGVGVNVELTDELDTVPCDFDGLTHGGYT  
APQRDPHTGELHAVSYSFARGHRVQSVITGDGHARTVDIEVAGSPMMHSFSLTN  
YVYIDLPVTFDPMPVPSVPRWLQRPALVIOVLGRVRIPOD IAALGNRMQGHSD  
RLPVAWNPSPARVGVMPREGNEDVRWEDIEPCYVYHPLNAYSECRNCAEVLVDYV  
RYSRMFDRDRPGCDSPSLDRWTINLATCAVTAECRDRQAQEFPRINELVGGPHR  
FAYTVIGEGFLVGAGAAALSTPLYKQDCVTGSSVSLDPLLLIGEMVFPVNPNSARAE  
DQGLIMGVYWHGRDEGQLLLDQAOTLESIA TVLHPQRPVPMGFHGNAPTt"  
1585..2664

## gene

## CDS

/gene="Rv0655"  
1585..2664  
/gene="Rv0655"  
FASTA score: YRBF\_ECOLI\_P45393 hypothetical abc  
transporter atp-binding (269 aa) opt: 644 z-score: 721.8  
E(): 3.4e-33 (38.5% identity in 244 aa overlap); contains  
PS00017 ATP/GTP-binding site motif A, PS00211 ABC  
transporters family signature, highly similar to M. leprae  
MKL\_MYCLE\_P30769 possible ribonucleotide transport atp-  
(347 aa) opt: 2021 z-score:2244.4 E(): 0. (92.2% identity  
in 335 aa overlap). Also similar to many otherM.  
tuberculosis ABC transporters eg. MTCY253.24 (33.6%  
identity in 241 aa overlap)"  
/codon\_start=1  
/transl\_table=11  
/product="hypothetical protein Rv0655"  
/protein\_id="CAB09379.1"  
/db\_xref="GI:2143306"  
/db\_xref="SWISS-PROT:O06784"  
/translation="MRYSDSYHTTGRMQPRASTEGFPMGVSIEVNLTKSFGSSRIWE  
DVTITPAGEVSVLLGPGSGTKSVLKLIGLLRPERGSIIDGTIIIECSAKELYEI  
RTLFGVLFDQDGLFGSMNLVYNTAPFLREHTKKVESEIRDIVMEKLVGLGCGDEKFK  
PGEISGGMKRGAGLARALVDPQILCDEPDGSLDPVRTAVLSOLINDINAQIDATIL  
IVTNINARTVPDNGMLFRKHLVMEGPREVILTSDEPVVRFQNLNGRRIGPICMSEE  
KDEATMAEEQALLDAGHHAGGVEIEGVPPQISATPGMPERRKAVARQARVREMLHTL  
PKKAAATLDDLEGTHKAYVHEIGQ"  
1762..1785  
/gene="Rv0655"  
/note="PS00017 ATP/GTP-binding site motif A"  
2074..2118  
/gene="Rv0655"  
/note="PS00211 ABC transporters family signature"  
complement(3052..3435)  
/gene="Rv0656c"  
complement(3052..3435)  
/gene="Rv0656c"  
/note="Rv0656c, (MTC1376.20, unknown), len: 127 aa"  
/codon\_start=1  
/transl\_table=11  
/product="hypothetical protein Rv0656c"  
/protein\_id="CAB09378.1"  
/db\_xref="GI:2143305"  
/db\_xref="SPTREMBL:O06783"  
/translation="MAAATTTGTHRGLELRAAQRAVGSCEPQRAEFCRSANNADEFDO  
MSRMFGDVPVDPVPKPSVWRWIDSAQHLRAGAVGALSVVVDLLICDTAAARGLVVLH  
DDADYELAERHLPDIRVRVVSADD"  
complement(3530..3685)  
/gene="Rv0657c"  
complement(3530..3685)  
/note="Rv0657c"  
/note="Rv0657c, (MTC1376.19), unknown, len: 51 aa; similar  
to several other M. tuberculosis hypothetical proteins eg.  
YW08\_MYCTU\_Q10848 hypothetical 8.9 kd protein cy39.08c (80  
aa) FASTA scores: opt: 107 z-score: 182.3 E(): 0.0038,  
45.8% identity in 48 aa overlap. Also similar to MTCY48\_5  
andAL020958|SC4H8.7 Streptomyces coelicolor cosmid 4H8 (66  
aa), 41.0% identity in 39 aa overlap."  
/codon\_start=1  
/transl\_table=11  
/product="hypothetical protein Rv0657c"  
/protein\_id="CAB09377.1"  
/db\_xref="GI:2143304"  
/db\_xref="SPTREMBL:O06782"  
/translation="MSVTQIDLDEALADVMRIAATVHTKKEAVNLAMRDYVERFRRIE  
ALASRR"  
complement(3691..3696)  
/note="possible RBS upstream of Rv0657c"  
complement(3761..4477)  
/gene="Rv0658c"  
complement(3761..4477)  
/gene="Rv0658c"  
/note="Rv0658c, (MTC1376.18), len: 238, unknown,

probablemembrane protein, similar to YPRB\_ECOLI P33774  
hypothetical24.3 kd protein (urf 1) (217 aa), fasta  
scores: opt: 174 z-score: 215.8 E(1): 5.3e-05, (25.6%  
identity in 223 aa overlap). Also similar to MTCY359.10  
(28.7% identity in 178 aa overlap)  
/codon\_start=1  
/transl\_table=11  
/product="hypothetical protein Rv0658c"  
/protein\_id="CAB09376.1"  
/db\_xref="GI:2143303"  
/db\_xref="SPTREMBL:O06781"  
/translation="MEAGADTVAPSHRWGLGAFVLVFLVASTSLAVLTGCHGV  
SAGVIALALARTVVAAGLALITRLKNGNLTDLRLKWSHRLGLMFGFGMLVT  
IPASLVTAIVGPGNSAVRIFGVRASWPALVFLVVFVAPLCEIIIRGLLWG  
IVDRWGMALVTVTVFALAHLEFARAPLLVVAIPALAREFVSGLLASIVTHQV  
TNLLPGIVLLGLTGATSLP"  
/note="possible RBS upstream of Rv0658c"  
/complement(4480..4483)  
/note="possible RBS upstream of Rv0658c"  
/complement(4753..5061)  
/gene="Rv0659c"  
/complement(4753..5061)  
/gene="Rv0659c"  
/complement(4753..5061)  
/note="Rv0659c, (MTCI376.17), len: 102; unknown, similar  
to YW28\_MYCTU Q10867 hypothetical 12.3 kd protein cy39.28  
(114 aa), fasta scores: opt: 144 z-score: 213.2 E(1):  
7.3e-05, 30.8% identity in 107 aa overlap. Also similar to  
MTCY099.22 (32.7% identity in 101 aa overlap)"  
/codon\_start=1  
/transl\_table=11  
/product="hypothetical protein Rv0659c"  
/protein\_id="CAB09387.1"  
/db\_xref="GI:2143302"  
/db\_xref="SPTREMBL:O06780"  
/translation="MRGELMFAATPGGDRVLVLTDRPVDVADRVVVALTRIRG  
LVSELELFAVENRPSDCVVFNDIHTLPTAFRRIRLSPARLHEACQTLRASTGC  
"  
/complement(5048..5293)  
/gene="Rv0660c"  
/complement(5048..5293)  
/note="Rv0660c"  
/note="Rv0660c, (MTCI376.16), len: 81, some similarity to  
IAF016485\_130 Halobacterium sp: NRC-1 plasm (100 aa),  
32.4% identity in 74 aa overlap"  
/codon\_start=1  
Query Match 55.5%; Score 123.8; DB 1; Length 19770;  
Best Local Similarity 76.7%; Pred. No. 6e-14;  
Matches 171; Conservative 0; Mismatches 37; Indels 15; Gaps 1;  
QY 1 TCAAGGAGAGCGCTACGACCTGGCCCGCTCGCGCTACAGGTCAACAAGAAGCTGG 60  
DB 10758 TCAAGGAGAGCGCTACGACCTGGCCCGCTCGCGCTATAGGTCAACAAGAAGCTGG 10817  
QY 61 GCTCTGAACACCGATCATCCGATCACCAACGACGCTGACCGAAGAAGACGCTGTCGCCA 120  
DB 10818 GCGTGCATGTCGCGGAGCCATCACGCTCGACGCTGACCGAAGAAGACGCTGTCGCCA 10877  
QY 121 CCATCCAGTACCTGGTTCCTCCGACCAACGCTCTCAGGTCGCCAGGCCCGCTTATGA 180  
DB 10878 CCATCCAGTACCTGGTTCCTCCGACCAACGCTCTCAGGTCGCCAGGCCCGCTTATGA 180  
QY 181 CTGTCCCGCGGGTTCGAGGTGCCGGTGGAAACGACGACAT 223  
DB 10923 CCGTTCGGCGCGCTGAGGTGCCGGTGGAAACGACGACAT 10965  
RESULT 15  
MSU24494  
LOCUS MSU24494 3752 bp DNA linear BCT 02-MAR-2000  
DEFINITION Mycobacterium smegmatis DNA polymerase (rpoB) gene, complete cds.  
ACCESSION U24494  
VERSION U24494.1 GI:790347  
KEYWORDS

SOURCE  
ORGANISM Mycobacterium smegmatis  
Mycobacterium smegmatis  
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
Actinomycetales; Corynebacterineae; Mycobacteriaceae;  
Mycobacterium.  
REFERENCE 1 (bases 1 to 3752)  
AUTHORS Hetherington, S.V., Watson, A.S. and Patrick, C.C.  
TITLE Sequence and analysis of the rpoB gene of Mycobacterium smegmatis  
JOURNAL Antimicrob. Agents Chemother. 39 (9), 2164-2166 (1995)  
MEDLINE 96050766  
PUBMED 8540740  
REFERENCE 2 (bases 1 to 3752)  
AUTHORS Hetherington, S.V.  
TITLE Direct Submission  
JOURNAL Submitted (11-APR-1995) Seth V. Hetherington, Infectious Diseases,  
St. Jude Children's Research Hospital, 332 N. Lauderdale, Memphis,  
TN 38101, USA  
FEATURES  
Location/Qualifiers  
source 1..3752  
/organism="Mycobacterium smegmatis"  
/db\_xref="taxon:1772"  
gene 194..3703  
/gene="rpoB"  
CDS 194..3703  
/gene="rpoB"  
/codon\_start=1  
/evidence=experimental  
/transl\_table=11  
/product="DNA polymerase"  
/protein\_id="AA091426.2"  
/db\_xref="GI:7144498"  
/translation="MLEGCIIVASSQSKNAITNNSVRGAPNGVSPAKLREPLEVRGL  
LDVQDSFEVLGVRVRRQAAIDRGEPVGLLEEVLAELSPIEDFSGMSLSFSDPR  
FDEYKASVDCKDKMTYAAPLFTVAETFINNTGEIKSOTVFMGDFPMTEKGTFIIN  
GTERVYQSLVRSFGVYFDETIDKSTKTLHSVKVIPCGAWLEFDFDKRDTVGVRIID  
RKRPQVTVLLKALGVFNEOIVVERFGSEIMMGTLEKDTTSGTDEALDLYLRKLPCE  
PPTKESAOITLLENLFFKEXBYDLARVORYKVKKGLGNACKPITSSLTTEEDVATIE  
YLVRHGGQTSVTPVGGVEFAVESHDIHFGENRRLRTVYGGELIQNOIKRVLRSRERVVR  
ERMTTQDVEAITPOTLINIRPVAAIKIEFGTSQSFQMDONNPLSGLTHKRRLSALG  
PGLSRERAGLEVDRVHPSHYGRMCPIETPEGNIGLISLSVYARVNGFETPYR  
KVGVVYTDQIDYLTADEEDRHVAQNSPTDEGRFTEDRVVVRKKGGEVFVSADQ  
VDYMDVSPROMVSATAMIPLEHDDANRALMGANMORQAVPLVRSEAPLVGTGMELR  
AADVATWSSHKTVIEVSADYITVMADDTQSYRLKRFARSHGCTCANORPVIDA  
GORVEAGQVADGPTONGEMALGKNLLVAIMPWEHNYEDAILNSRLVEEDVLTSL  
HIEHEIDARDTKLGAEEITRVIPKLSDEVLADLDEPRIVRIGAEVRDGLVGVKTP  
KGTELTPEERLLRAIFEGKAREVRDTSKRVPHGESKVIIGRVFSDDEDELPLGVN  
ELRVVYVAQRKISDQKLLARRHGNKVGIGKILPVEDMPFLPDGTPVDIILNTHGCRV  
VLNTQILETHLGMWAGKAGWNIDVLACVPDWSKLPPELYSAPADSVATVPFDGAGE  
GELAGLGGSTLPNRDGEVMVNDCKATLFDGRSSGPEFPYPTVTVGYMYILKLHLVDD  
KIHARSTGPYSMTIQQOPLGCKAOGGORGEMSCWAMQAVCAAYTLOELLTIKSDPTV  
GRVYVYAIKGENIPPEGPIDSEFKVLLKELQSLCLNVEVLSDDGRAIENRDGDEDL  
ERAAANGLINLSRNESASVEDLA"  
BASE COUNT 706 a 1225 c 1210 g 611 t  
ORIGIN  
Query Match 55.1%; Score 122.8; DB 1; Length 3752;  
Best Local Similarity 88.7%; Pred. No. 1.1e-13;  
Matches 133; Conservative 0; Mismatches 17; Indels 0; Gaps 0;  
QY 1 TCAAGGAGAGCGCTACGACCTGGCCCGCTCGCGCTACAAAGTCAACAAGAAGCTGG 60  
DB 1068 TCAAGGAGAGCGCTACGACCTGGCCCGCTGGCCGTTTACAAAGTCAACAAGAAGCTGG 1127  
QY 61 GCCTGAACACCGATCATCCGATCACCAACGACGCTGACCGAAGAAGACGCTGTCGCCA 120  
DB 1128 GCCTGAACCGGGAAGCGGATCATCCAGCTCGACGCTGACCGAAGAAGACGCTGTCGCCA 1187  
QY 121 CCATCCAGTACCTGGTTCCTCCGACCAACGACGACAT 150  
DB 1188 CCATCCAGTACCTGGTTCCTCCGACGAGG 1217  
Search completed: November 13, 2002, 01:26:23

Job time : 709.59 secs

---

GenCore version 5.1.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 12, 2002, 15:44:35 : Search time 146.046 Seconds  
(without alignments)  
3438.621 Million cell updates/sec

Title: US-09-697-123B-13

Perfect score: 223

Sequence: 1 tcaaggaggaagcgctacgac.....ccgtggaaccgacgacat 223

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :				N_Geneseq_101002:*	
	1:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:*			
	2:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:*			
	3:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:*			
	4:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:*			
	5:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:*			
	6:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:*			
	7:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:*			
	8:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:*			
	9:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:*			
	10:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:*			
	11:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:*			
	12:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:*			
	13:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:*			
	14:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:*			
	15:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:*			
	16:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:*			
	17:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:*			
	18:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:*			
	19:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:*			
	20:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:*			
	21:	/SID52/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:*			
	22:	/SID52/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:*			
	23:	/SID52/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:*			
	24:	/SID52/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:*			

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	223	100.0	223	AAS05213	Mycobacterium gast
2	178.2	79.9	223	AAS05209	Mycobacterium scro
3	152.6	68.4	208	AAS05208	Mycobacterium kans
4	141.4	63.4	208	AAS05215	Mycobacterium aviu
5	138.2	62.0	208	AAS05217	Mycobacterium cels
6	135	60.5	208	AAS05201	Mycobacterium gord
7	135	60.5	208	AAS05219	Mycobacterium fort
8	135	60.5	208	AAS05224	Mycobacterium xeno
9	133.4	59.8	208	AAS05203	Mycobacterium gord

10	133.4	59.8	208	22	AAS05211	Mycobacterium mari
11	131.8	59.1	208	22	AAS05202	Mycobacterium gord
12	131.8	59.1	208	22	AAS05210	Mycobacterium ulce
13	126.6	56.8	207	22	AAS05204	Mycobacterium gord
14	123.8	55.5	208	22	AAS05205	Mycobacterium tube
15	123.8	55.5	208	22	AAS05206	Mycobacterium terr
16	123.8	55.5	208	22	AAS05216	Mycobacterium bovi
17	123.8	55.5	970	17	AAT09676	Mycobacterium tube
18	123.8	55.5	3519	22	AAH51976	Mycobacterium tube
19	123.8	55.5	3534	22	AAH02079	Mycobacterium tube
20	123.8	55.5	3853	21	AAA74651	M. tuberculosis rp
21	123.8	55.5	3853	21	AA889994	Mycobacterium afri
22	122.2	54.8	208	22	AAS05222	Mycobacterium intr
23	121	54.3	205	22	AAS05220	Mycobacterium haem
24	120.6	54.1	208	22	AAS05223	Mycobacterium szul
25	118.2	53.0	207	22	AAS05212	M.leprae rpoB gene
26	111	49.8	3447	14	AAQ51532	Mycobacterium abs
27	107.8	48.3	214	22	AAS05221	Mycobacterium flav
28	104.4	46.8	211	22	AAS05218	Mycobacterium malm
29	101.4	45.5	214	22	AAS05214	Mycobacterium chel
30	99.8	44.8	214	22	AAS05207	C glutamicum codin
31	62.8	28.2	3495	22	AAH65512	C glutamicum codin
32	62.8	28.2	34980	22	AAH68525	Propionibacterium
33	52	23.3	27426	23	AAS59541	S. chrysomallus ac
34	42	18.8	3849	22	AAF25795	A. chrysogenum pho
35	41.2	18.5	3306	13	AAQ22202	Phosphoglycerate k
36	41.2	18.5	3306	13	AAQ23005	PGK. Acremonium c
37	41.2	18.5	3306	14	AAQ48534	Nucleotide sequenc
38	41.2	18.5	77536	21	AAA14651	Corr tassal-derive
39	41	18.4	278	24	ABL72018	Pseudomonas aerugi
40	39	17.5	471	23	AAS54080	Nucleotide sequenc
41	39	17.5	5857	21	AAA58471	Superoxide dismuta
42	38.8	17.4	491	15	AAQ62119	Maize nitrite redu
43	37.8	17.0	1844	12	AAQ12000	Pseudomonas aerugi
44	37.2	16.7	1350	23	AAS54328	P. aeruginosa accB
45	37.2	16.7	2760	16	AAQ79921	

ALIGNMENTS

RESULT 1

AAS05213

ID AAS05213 standard; DNA; 223 BP.

AC AAS05213:

XX 07-SEP-2001 (first entry)

DT Mycobacterium gastri rpoB gene fragment.

DE Non-tuberculous mycobacteria: rpoB gene fragment; NTM; HIV; PRA; RFLP;

XX PCR-restriction fragment length polymorphism analysis; ds.

OS Mycobacterium gastri.

XX WO200131061-A1.

XX 03-MAY-2001.

XX 27-OCT-2000; 2000WO-KR01223.

XX 27-OCT-1999; 99KR-0046795.

XX (ERUM-) ERUME BIOTECH CO LTD.

XX Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;

XX WPI: 2001-300520/31.

XX New DNA fragments from the rpoB gene of mycobacteria, useful for

PT diagnosis and identification of many mycobacterial species by

PT restriction fragment length polymorphism



```
XX PS Claim 1; Page 44; 50pp; English.
XX CC The present sequence for Mycobacterium gastri rpoB gene
CC CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
CC CC various Mycobacterial species. These rpoB gene fragments can be used
CC CC in the diagnosis and identification of Mycobacterium species using a
CC CC novel PCR-restriction fragment length polymorphism analysis (PRA)
CC CC method. The method comprises obtaining a restriction fragment length
CC CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
CC CC amplifying and digesting the DNA fragment from the known rpoB gene
CC CC fragments with the unidentified fragment. The rpoB gene fragments
CC CC are useful to identify a wide range of Mycobacterium species, e.g. for
CC CC diagnosis or to obtain epidemiological and pathogenesis information for
CC CC selection of appropriate therapies, including M. tuberculosis, M. leprae
CC CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene
CC CC fragments is rapid, precise, simple and cost effective (only 1 PCR
CC CC required), and can differentiate between many species in a single
CC CC experiment, including those difficult to distinguish by usual biochemical
CC CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
CC CC detecting specific Mycobacterial species.
XX SQ Sequence 223 BP; 49 A; 77 C; 65 G; 32 T; 0 other;

Query Match 100.0%; Score 223; DB 22; Length 223;
Best Local Similarity 100.0%; Pred. No. 9,7e-44;
Matches 223; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCAAGGAGAGCGGTACGACCTGGCCCGCTCGCGCGCTACAAAGGTCAACAAGAGCTGG 60
DB 1 TCAAGGAGAGCGGTACGACCTGGCCCGCTCGCGCGCTACAAAGGTCAACAAGAGCTGG 60

QY 61 GCCTCAACACCGGATCCGATACACACGACGCTGACCGAAGAGAGCTGTCGCCCA 120
DB 61 GCCTCAACACCGGATCCGATACACACGACGCTGACCGAAGAGAGCTGTCGCCCA 120

QY 121 CCATCGAGTACCTGTTGGCTGCGCCACGCGCTCTCAGGTGGCCAGGCGCCGTTATGA 180
DB 121 CCATCGAGTACCTGTTGGCTGCGCCACGCGCTCTCAGGTGGCCAGGCGCCGTTATGA 180

QY 181 CTGTCCCGCGGGTCCAGTCCGCTGCGGTGGAACCGACGACAT 223
DB 181 CTGTCCCGCGGGTCCAGTCCGCTGCGGTGGAACCGACGACAT 223

RESULT 2
AAS05209
ID AAS05209 standard; DNA; 223 BP.
XX AC AAS05209;
XX DT 07-SEP-2001 (first entry)
XX DE Mycobacterium scrofulaceum rpoB gene fragment.
XX KW Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;
XX KW PCR-restriction fragment length polymorphism analysis; ds.
XX OS Mycobacterium scrofulaceum.
XX PN WO200131061-A1.
XX PD 03-MAY-2001.
XX PF 27-OCT-2000; 2000WO-KR01223.
XX PR 27-OCT-1999; 99KR-0046795.
XX PA (ERUM-) ERUME BIOTECH CO LTD.
XX PJ Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;
```

```
XX DR WPI; 2001-300520/31.
XX PT New DNA fragments from the rpoB gene of mycobacteria, useful for
XX PT diagnosis and identification of many mycobacterial species by
XX PS restriction fragment length polymorphism .
XX PS Claim 1; Page 42-43; 50pp; English.
XX CC The present sequence for Mycobacterium scrofulaceum rpoB gene
XX CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
XX CC various Mycobacterial species. These rpoB gene fragments can be used
XX CC in the diagnosis and identification of Mycobacterium species using a
XX CC novel PCR-restriction fragment length polymorphism analysis (PRA)
XX CC method. The method comprises obtaining a restriction fragment length
XX CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
XX CC amplifying and digesting the DNA fragment from the microorganism to
XX CC be identified and comparing the RFLP patterns from the known rpoB gene
XX CC fragments with the unidentified fragment. The rpoB gene fragments
XX CC are useful to identify a wide range of Mycobacterium species, e.g. for
XX CC diagnosis or to obtain epidemiological and pathogenesis information for
XX CC selection of appropriate therapies, including M. tuberculosis, M. leprae
XX CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene
XX CC fragments is rapid, precise, simple and cost effective (only 1 PCR
XX CC required), and can differentiate between many species in a single
XX CC experiment, including those difficult to distinguish by usual biochemical
XX CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
XX CC detecting specific Mycobacterial species.
XX SQ Sequence 223 BP; 45 A; 77 C; 71 G; 30 T; 0 other;

Query Match 79.9%; Score 178.2; DB 22; Length 223;
Best Local Similarity 87.4%; Pred. No. 3.5e-33;
Matches 195; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

QY 1 TCAAGGAGAGCGGTACGACCTGGCCCGCTCGCGCGCTACAAAGGTCAACAAGAGCTGG 60
DB 1 TCAAGGAGAGCGGTACGACCTGGCCCGCTCGCGCGCTACAAAGGTCAACAAGAGCTGG 60

QY 61 GCCTCAACACCGGATCCGATACACACGACGCTGACCGAAGAGAGCTGTCGCCCA 120
DB 61 GTCTGCGCGCGGGGAGCGGATCAGCTGCTCAGCTGACCGGAGGAGAGCTGTCGCCGA 120

QY 121 CCATCGAGTACCTGTTGGCTGCGCTGACCGACGCTCTCAGGTGGCCAGGCGCCGTTATGA 180
DB 121 CCATCGAATACCTGTTGGCTGCGCTGACCGACGCTCTCAGGTGGCCAGGCGCGCTCATGA 180

QY 181 CTGTCCCGCGGGTCCAGTCCGCTGCGGTGGAACCGACGACAT 223
DB 181 CTGTCCCGCGGGTCCAGTCCGCTGCGGTGGAACCGACGACAT 223

RESULT 3
AAS05208
ID AAS05208 standard; DNA; 208 BP.
XX AC AAS05208;
XX DT 07-SEP-2001 (first entry)
XX DE Mycobacterium kansasii rpoB gene fragment.
XX KW Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;
XX KW PCR-restriction fragment length polymorphism analysis; ds.
XX OS Mycobacterium kansasii.
XX PN WO200131061-A1.
XX PD 03-MAY-2001.
XX PF 27-OCT-2000; 2000WO-KR01223.
```

```

XX 27-OCT-1999; 99KR-0046795.
XX (ERUM-) ERUME BLOPFCH CO LTD.
XX Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ.
XX WPI: 2001-300520/31.
XX New DNA fragments from the rpoB gene of mycobacteria, useful for
XX diagnosis and identification of many mycobacterial species by
XX restriction fragment length polymorphism -
XX Claim 1: Page 42; 50pp: English.
XX The present sequence for Mycobacterium kansasii rpoB gene
XX fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
XX various Mycobacterial species. These rpoB gene fragments can be used
XX in the diagnosis and identification of Mycobacterium species using a
XX novel PCR-restriction fragment length polymorphism analysis (PRA)
XX method. The method comprises obtaining a restriction fragment length
XX polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
XX amplifying and digesting the DNA fragment from the microorganism to
XX be identified and comparing the RFLP patterns from the known rpoB gene
XX fragments with the unidentified fragment. The rpoB gene fragments
XX are useful to identify a wide range of Mycobacterium species, e.g. for
XX diagnosis or to obtain epidemiological and pathogenesis information for
XX selection of appropriate therapies, including M. tuberculosis, M. leprae
XX and non-tuberculous mycobacteria (NTM) encountered in subjects infected
XX with human immunodeficiency virus (HIV). Analysis of the rpoB gene
XX fragments is rapid, precise, simple and cost effective (only 1 PCR
XX required), and can differentiate between many species in a single
XX experiment, including those difficult to distinguish by usual biochemical
XX tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
XX detecting specific Mycobacterial species.
XX Sequence 208 BP; 51 A; 65 C; 65 G; 27 T; 0 other;
SQ
Query Match 68.4%; Score 152.6; DB 22; Length 208;
Best Local Similarity 84.8%; Pred. No. 3.8e-27;
Matches 189; Conservative 0; Mismatches 19; Indels 15; Gaps 1;
QY 1 TCAAGGAGAGCGCTACGACCTGGCGCGCTGCGCGCTACAAAGTCAACAAGAAGCTGG 60
DB 1 TCAAGGAGAGCGCTACGACCTGGCGCGCTGTCGGCGGATACAAAGTCAACAAGAAGCTGG 60
QY 61 GCCTGAACACCGATCATCGGATCACACACAGCGCTGACCGAAGAACGCTCGTCGCCA 120
DB 61 GCCTGAACACCAATCATCGGATCACACAGCGACGCTGACCGAAGAACGCTCGTCGCCA 120
QY 121 CCATCGAGTACCTGGTTCGCCCTGACCGACGCGCTCTCAGGGTGGCCAGGCCCGCTTATGA 180
DB 121 CCATCGAGTATCTGGTTCGCCCTGACAGAGGGC-----CAGGCCAGCATGA 165
QY 181 CTGTCGCCCGGGGTGAGAGGTGCGCGGTGGAAACCGACCATAT 223
DB 166 CCGTGCCTGGCGGGGTGAGAGGTGCGCGGTGGAACCGACCATAT 208
RESULT 4
AAS05215
ID AAS05215 standard; DNA: 208 BP.
XX AAS05215;
XX 07-SEP-2001 (first entry)
XX Mycobacterium avium rpoB gene fragment.
XX Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;
XX PCR-restriction fragment length polymorphism analysis; ds.
XX Mycobacterium avium.

```

```

XX WO200131061-A1.
XX 03-MAY-2001.
XX 27-OCT-2000; 2000WO-KR01223.
XX 27-OCT-1999; 99KR-0046795.
XX (ERUM-) ERUME BIOTECH CO LTD.
XX Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ.
XX WPI: 2001-300520/31.
XX New DNA fragments from the rpoB gene of mycobacteria, useful for
XX diagnosis and identification of many mycobacterial species by
XX restriction fragment length polymorphism -
XX Claim 1: Page 44; 50pp: English.
XX The present sequence for Mycobacterium avium rpoB gene
XX fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
XX various Mycobacterial species. These rpoB gene fragments can be used
XX in the diagnosis and identification of Mycobacterium species using a
XX novel PCR-restriction fragment length polymorphism analysis (PRA)
XX method. The method comprises obtaining a restriction fragment length
XX polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
XX amplifying and digesting the DNA fragment from the microorganism to
XX be identified and comparing the RFLP patterns from the known rpoB gene
XX fragments with the unidentified fragment. The rpoB gene fragments
XX are useful to identify a wide range of Mycobacterium species, e.g. for
XX diagnosis or to obtain epidemiological and pathogenesis information for
XX selection of appropriate therapies, including M. tuberculosis, M. leprae
XX and non-tuberculous mycobacteria (NTM) encountered in subjects infected
XX with human immunodeficiency virus (HIV). Analysis of the rpoB gene
XX fragments is rapid, precise, simple and cost effective (only 1 PCR
XX required), and can differentiate between many species in a single
XX experiment, including those difficult to distinguish by usual biochemical
XX tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
XX detecting specific Mycobacterial species.
XX Sequence 208 BP; 44 A; 69 C; 69 G; 26 T; 0 other;
SQ
Query Match 63.4%; Score 141.4; DB 22; Length 208;
Best Local Similarity 81.6%; Pred. No. 1.7e-24;
Matches 192; Conservative 0; Mismatches 26; Indels 15; Gaps 1;
QY 1 TCAAGGAGAGCGCTACGACCTGGCGCGCTGCGCGCTACAAAGTCAACAAGAAGCTGG 60
DB 1 TCAAGGAGAGCGCTACGACCTGGCGCGCTGCGCGCTACAAAGTCAACAAGAAGCTGG 60
QY 61 GCCTGAACACCGATCATCGGATCACACACAGCGCTGACCGAAGAACGCTCGTCGCCA 120
DB 61 GCCTGAACACCGGCTGAGCGGATCACAGGCTGACCGAAGAACGCTCGTCGCCA 120
QY 121 CCATCGAGTACCTGGTTCGCCCTGACCGACGCGCTCTCAGGGTGGCCAGGCCCGCTTATGA 180
DB 121 CCATCGAGTACCTGGTTCGCCCTGACAGAGG-----GTGAGGCCAGCATGA 165
QY 181 CTGTCGCCCGGGGTGAGAGGTGCGCGGTGGAAACCGACCATAT 223
DB 166 CCGTGCCTGGCGGGGTGAGAGGTGCGCGGTGGAACCGACCATAT 208
RESULT 5
AAS05217
ID AAS05217 standard; DNA: 208 BP.
XX AAS05217;
XX 07-SEP-2001 (first entry)
XX

```





```
Db 1 TCAAGGAGAAAGCGCTACGACCTGGCCGCGTGGCCGCTACAAGGTCAACAAGAACTCG 60
QY 61 GCCTGAACACCGATATCCGATCACCACACGACGCTGACCGAAGAGAGGTGTCGCCCA 120
Db 61 GCCTGAACACCGAATGCGCCAAACACACGACCTGACCGAAGAGAGGTGTCGCCCA 120
QY 121 CCATCGAGTACCTGGTTCGGCTCGACACACGCTCTCAGGGTGGCCAGGCCCGCTTATGA 180
Db 121 CCATCGAATACCTGGTTCGGCTTCACGAGG- - - - -GGCAGCGCCACGATGA 165
QY 181 CTGTCCCGGGGGTTCGAGTGGCGGTGGAACCGACGACAT 223
Db 166 AGTCCCGGTGGCGTTCGAGTGGCGGTGGAGACCGACGACAT 208

RESULT 9
AAS05203
ID AAS05203 standard; DNA; 208 BP.
XX
AC AAS05203;
XX
DT 07-SEP-2001 (first entry)
XX
DE Mycobacterium gordonae type III rpoB gene fragment.
XX
KW Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;
KW PCR-restriction fragment length polymorphism analysis; ds.
XX
OS Mycobacterium gordonae type III.
XX
PN W0200131061-A1.
XX
PD 03-MAY-2001.
XX
PF 27-OCT-2000; 2000WO-KR01223.
XX
PK 27-OCT-1999; 99KR-0046795.
XX
PA (ERUM-) ERUME BIOTECH CO LTD.
XX
PI Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;
XX
DR WPI; 2001-300520/31.
XX
PS Claim 1; Page 41; 50pp; English.
XX
CC The present sequence for Mycobacterium gordonae type III rpoB gene
CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
CC various Mycobacterial species. These rpoB gene fragments can be used
CC in the diagnosis and identification of Mycobacterium species using a
CC novel PCR-restriction fragment length polymorphism analysis (PRA)
CC method. The method comprises obtaining a restriction fragment length
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
CC amplifying and digesting the DNA fragment from the microorganism to
CC be identified and comparing the RFLP patterns from the known rpoB gene
CC fragments with the unidentified fragment. The rpoB gene fragments
CC are useful to identify a wide range of Mycobacterium species, e.g. for
CC diagnosis or to obtain epidemiological and pathogenesis information for
CC selection of appropriate therapies, including M. tuberculosis, M. leprae
CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene
CC fragments is rapid, precise, simple and cost effective (only 1 PCR
CC required), and can differentiate between many species in a single
CC experiment, including those difficult to distinguish by usual biochemical
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
CC detecting specific Mycobacterial species.
XX
SQ Sequence 208 BP; 46 A; 69 C; 63 G; 30 T; 0 other;
```

```
Query Match 59.8%; Score 133.4; DB 22; Length 208;
Best Local Similarity 79.4%; Pred. No. 1.3e-22;
Matches 177; Conservative 0; Mismatches 31; Indels 15; Gaps 1;
QY 1 TCAAGGAGAAAGCGCTACGACCTGGCCGCGTGGCCGCTACAAGGTCAACAAGAACTCG 60
Db 1 TCAAGGAGAAAGCGCTACGACCTGGCCGCGTGGCCGCTACAAGGTCAACAAGAACTCG 60
QY 61 GCCTGAACACCGATATCCGATCACCACACGACGCTGACCGAAGAGAGGTGTCGCCCA 120
Db 61 GCCTGACACGTCGGCGATCCGATCACCAGCTCCAGCTGACCGAAGAGAGGTGTCGCCCA 120
QY 121 CCATCGAGTACCTGGTTCGGCTCGACACGCTCTCAGGGTGGCCAGGCCCGCTTATGA 180
Db 121 CCATCGAGTACCTGGTTCGGCTCGACGAGGTCAGCA- - - - -CACGATGA 165
QY 181 CTGTCCCGGGGGTTCGAGTGGCGGTGGAACCGACGACAT 223
Db 166 CCGTCCGGCGCACCGAGGTTCGCGTGGAGACCGACGACAT 208

RESULT 10
AAS05211
ID AAS05211 standard; DNA; 208 BP.
XX
AC AAS05211;
XX
DT 07-SEP-2001 (first entry)
XX
DE Mycobacterium marinum rpoB gene fragment.
XX
KW Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;
KW PCR-restriction fragment length polymorphism analysis; ds.
XX
OS Mycobacterium marinum.
XX
PN W0200131061-A1.
XX
PD 03-MAY-2001.
XX
PF 27-OCT-2000; 2000WO-KR01223.
XX
PK 27-OCT-1999; 99KR-0046795.
XX
PA (ERUM-) ERUME BIOTECH CO LTD.
XX
PI Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;
XX
DR WPI; 2001-300520/31.
XX
PS Claim 1; Page 43; 50pp; English.
XX
CC The present sequence for Mycobacterium marinum rpoB gene
CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
CC various Mycobacterial species. These rpoB gene fragments can be used
CC in the diagnosis and identification of Mycobacterium species using a
CC novel PCR-restriction fragment length polymorphism analysis (PRA)
CC method. The method comprises obtaining a restriction fragment length
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
CC amplifying and digesting the DNA fragment from the microorganism to
CC be identified and comparing the RFLP patterns from the known rpoB gene
CC fragments with the unidentified fragment. The rpoB gene fragments
CC are useful to identify a wide range of Mycobacterium species, e.g. for
CC diagnosis or to obtain epidemiological and pathogenesis information for
CC selection of appropriate therapies, including M. tuberculosis, M. leprae
CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene
CC fragments is rapid, precise, simple and cost effective (only 1 PCR
```

CC required), and can differentiate between many species in a single  
CC experiment, including those difficult to distinguish by usual biochemical  
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for  
CC detecting specific Mycobacterial species.

XX  
SQ Sequence 208 BP; 45 A; 69 C; 68 G; 26 T; 0 other;

Query Match 59.8%; Score 133.4; DB 22; Length 208;  
Best Local Similarity 79.4%; Pred. No. 1.3e-22;  
Matches 177; Conservative 0; Mismatches 31; Indels 15; Gaps 1;

QY 1 TCAAGGAGAGCGCTACGACCTGGCCCGCTCGCGCGGTACAAAGTCAACGAAGCTGG 60

Db 1 TCAAGGAGAGCGCTACGACCTGGCCCGCTCGCGCGGTACAAAGTCAACGAAGCTGG 60

QY 61 GCCTGAACACCATCATCGATCACCAACACGACGCTCAGCGAAGACGCTGTCGCCA 120

Db 61 GCCTGAACACCGCCGACCCATCACCACTCGACGCTCAGCGAAGACGCTGTCGCCA 120

QY 121 CCATCGAGTACCTGGTTGCGCTGCACACGCGCTCTCAGGGTGGCCGCCCGCTTATGA 180

Db 121 CCATCGAATACCTGGTTCGCTTTCACGAGGGC-----CAGACCGCGATGA 165

QY 181 CTGTCCCGCGGGTTCGAGGTGCGGTGGAAACCGACGACAT 223

Db 166 CGCTCCGCGGGTTCGAGGTGCGGTGGAGACCGACGACAT 208

## RESULT 11

AAS05202  
ID AAS05202 standard; DNA; 208 BP.

XX  
AC AAS05202;

XX  
DT 07-SEP-2001 (first entry)

XX  
DE Mycobacterium gordonae type II rpoB gene fragment.

XX  
KW Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;  
PCR-restriction fragment length polymorphism analysis; ds.

XX  
OS Mycobacterium gordonae type II.

XX  
PN WO200131061-A1.

XX  
PD 03-MAY-2001.

XX  
PF 27-OCT-2000; 2000WO-KR01223.

XX  
PR 27-OCT-1999; 99KR-0046795.

XX  
PA (ERUM-) ERUME BIOTECH CO LTD.

PI Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;

XX  
DR WPI; 2001-300520/31.

XX  
PT New DNA fragments from the rpoB gene of mycobacteria, useful for  
PT diagnosis and identification of many mycobacterial species by  
PT restriction fragment length polymorphism -

XX  
PS Claim 1; Page 40; 50pp; English.

XX  
CC The present sequence for Mycobacterium gordonae type II rpoB gene  
CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05242) from  
CC various Mycobacterial species. These rpoB gene fragments can be used  
CC in the diagnosis and identification of Mycobacterium species using a  
CC novel PCR-restriction fragment length polymorphism analysis (PRA)  
CC method. The method comprises obtaining a restriction fragment length  
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,  
CC amplifying and digesting the DNA fragment from the microorganism to  
CC be identified and comparing the RFLP patterns from the known rpoB gene  
CC fragments with the unidentified fragment. The rpoB gene fragments

CC are useful to identify a wide range of Mycobacterium species, e.g. for  
CC diagnosis or to obtain epidemiological and pathogenesis information for  
CC selection of appropriate therapies, including M. tuberculosis, M. leprae  
CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected  
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene  
CC fragments is rapid, precise, simple and cost effective (only 1 PCR  
CC required), and can differentiate between many species in a single  
CC experiment, including those difficult to distinguish by usual biochemical  
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for  
CC detecting specific Mycobacterial species.

XX  
SQ Sequence 208 BP; 45 A; 62 C; 73 G; 28 T; 0 other;

Query Match 59.1%; Score 131.8; DB 22; Length 208;  
Best Local Similarity 78.9%; Pred. No. 3.1e-22;  
Matches 176; Conservative 0; Mismatches 32; Indels 15; Gaps 1;

QY 1 TCAAGGAGAGCGCTACGACCTGGCCCGCTCGCGCGGTACAAAGTCAACGAAGCTGG 60

Db 1 TCAAGGAGAGCGCTACGACCTGGCCCGGTGGCGCGGTACAAAGTCAACGAAGCTGG 60

QY 61 GCCTGAACACCATCATCGATCACCAACGACGCTGACCGAAGACGCTGTCGCCA 120

Db 61 GTCTGAACGCTCGGCAAGCCGATCACCACTCGACGCTGACCGAAGACGCTGTCGCCA 120

QY 121 CCATCGAGTACCTGGTTGCGCTGCACACGCGCTCTCAGGGTGGCCGCCCGCTTATGA 180

Db 121 CCATCGAGTACCTGGTTCGCGCTGACGAGG-----GTCACTCGCGGATGA 165

QY 181 CTGTCCCGCGGGTTCGAGGTGCGGTGGAAACCGACGACAT 223

Db 166 CGGTTCCCGCGGGTTCGAGGTGCGGTGGAGACCGGACGACAT 208

## RESULT 12

AAS05210

ID AAS05210 standard; DNA; 208 BP.

XX  
AC AAS05210;

XX  
DT 07-SEP-2001 (first entry)

XX  
DE Mycobacterium ulcerans rpoB gene fragment.

XX  
KW Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;  
PCR-restriction fragment length polymorphism analysis; ds.

XX  
OS Mycobacterium ulcerans.

XX  
PN WO200131061-A1.

XX  
PD 03-MAY-2001.

XX  
PF 27-OCT-2000; 2000WO-KR01223.

XX  
PR 27-OCT-1999; 99KR-0046795.

XX  
PA (ERUM-) ERUME BIOTECH CO LTD.

PI Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;

XX  
DR WPI; 2001-300520/31.

XX  
PT New DNA fragments from the rpoB gene of mycobacteria, useful for  
PT diagnosis and identification of many mycobacterial species by  
PT restriction fragment length polymorphism -

XX  
PS Claim 1; Page 43; 50pp; English.

XX  
CC The present sequence for Mycobacterium ulcerans rpoB gene  
CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05242) from  
CC various Mycobacterial species. These rpoB gene fragments can be used  
CC in the diagnosis and identification of Mycobacterium species using a

CC novel PCR-restriction fragment length polymorphism analysis (PRA)  
 CC method. The method comprises obtaining a restriction fragment length  
 CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,  
 CC amplifying and digesting the DNA fragment from the microorganism to  
 CC be identified and comparing the RFLP patterns from the known rpoB gene  
 CC fragments with the unidentified fragment. The rpoB gene fragments  
 CC are useful to identify a wide range of Mycobacterium species, e.g. for  
 CC diagnosis or to obtain epidemiological and pathogenesis information for  
 CC selection of appropriate therapies, including M. tuberculosis, M. leprae  
 CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected  
 CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene  
 CC fragments is rapid, precise, simple and cost effective (only 1 PCR  
 CC required), and can differentiate between many species in a single  
 CC experiment, including those difficult to distinguish by usual biochemical  
 CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for  
 CC detecting specific Mycobacterium species.

XX Sequence 208 BP; 45 A; 68 C; 67 G; 28 T; 0 other;

Query Match 59.1%; Score 131.8; DB 22; Length 208;  
 Best Local Similarity 78.9%; Pred. No. 3.1e-22;  
 Matches 176; Conservative 0; Mismatches 32; Indels 15; Gaps 1;

Qy 1 TCAAGGAGAACCGCTACGACCTGCGCCGCTCGCCGCTACAAGGTCAACAAGAGCTGG 60  
 |||||  
 Db 1 TCAAGGAGAACCGCTACGACCTGCGCTCGCCGCTACAAGGTCAACAAGAGCTGG 60  
 |||||  
 Qy 61 CCTGAACACCGCATCCGATCACCGACGACGCTGACCGAAGAAGACGCTCGCGCCA 120  
 |||||  
 Db 61 CCTGAACACCGCGCCGACCCATCACCGCTGACCGCTGACCGAGAGACGCTCGCGCCA 120  
 |||||  
 Qy 121 CCATCGAGTACCTGGTTCGCTGACCGACGCGCTCTCAGGGTGGCCAGCCGCCCTTATGA 180  
 |||||  
 Db 121 CCATCGAATACCTGTGCTCGCTTGCACGAGGCG-----CAGACCGCGATGA 165  
 |||||  
 Qy 181 CTGTCCCGCGGGGTCGAGCTGCGGTGGAACCGACGACAT 223  
 |||||  
 Db 166 CGCTCCCGGGCGGTGCTGAGGTGCGGTCGAGACCGACGACAT 208  
 |||||

RESULT 13

AAS05204

ID AAS05204 standard; DNA; 207 BP.

XX AAS05204;

DT 07-SEP-2001 (first entry)

DE Mycobacterium gordonae type IV rpoB gene fragment.

KW Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;  
 KW PCR-restriction fragment length polymorphism analysis; ds.

OS Mycobacterium gordonae type IV.

PN WO200131061-A1.

PD 03-MAY-2001.

XX 27-OCT-2000; 2000WO-KR01223.

XX 27-OCT-1999; 99KR-0046795.

XX (ERUM-) ERUME BIOTECH CO LTD.

XX Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;

XX WPI; 2001-300520/31.

PT New DNA fragments from the rpoB gene of mycobacteria, useful for  
 PT diagnosis and identification of many mycobacterial species by  
 PT restriction fragment length polymorphism

XX

PS Claim 1; Page 41; 50pp; English.

XX The present sequence for Mycobacterium gordonae type IV rpoB gene  
 CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from  
 CC various Mycobacterium species. These rpoB gene fragments can be used  
 CC in the diagnosis and identification of Mycobacterium species using a  
 CC novel PCR-restriction fragment length polymorphism analysis (PRA)  
 CC method. The method comprises obtaining a restriction fragment length  
 CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,  
 CC amplifying and digesting the DNA fragment from the microorganism to  
 CC be identified and comparing the RFLP patterns from the known rpoB gene  
 CC fragments with the unidentified fragment. The rpoB gene fragments  
 CC are useful to identify a wide range of Mycobacterium species, e.g. for  
 CC diagnosis or to obtain epidemiological and pathogenesis information for  
 CC selection of appropriate therapies, including M. tuberculosis, M. leprae  
 CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected  
 CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene  
 CC fragments is rapid, precise, simple and cost effective (only 1 PCR  
 CC required), and can differentiate between many species in a single  
 CC experiment, including those difficult to distinguish by usual biochemical  
 CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for  
 CC detecting specific Mycobacterium species.

XX Sequence 207 BP; 45 A; 66 C; 66 G; 30 T; 0 other;

Query Match 56.8%; Score 126.6; DB 22; Length 207;

Best Local Similarity 77.6%; Pred. No. 5.2e-21;

Matches 173; Conservative 0; Mismatches 34; Indels 16; Gaps 1;

Qy 1 TCAAGGAGAACCGCTACGACCTGCGCCGCTCGCCGCTACAAGGTCAACAAGAGCTGG 60  
 |||||  
 Db 1 TCAAGGAGAACCGCTACGACCTGCGCCGCTCGCCGCTACAAGGTCAACAAGAGCTGG 60  
 |||||

Qy 61 GCCTGAACACCGATCATCCGATCACCGACGACGCTGACCGAAGAAGACGCTCGCGCCA 120  
 |||||  
 Db 61 GCCTGCAATGTGCGCGATCCGATCACCGACGCTGACCGAAGAAGAGACGCTCGCGCCA 120  
 |||||

Qy 121 CCATCGAGTACCTGGTTCGCTGACCGACGCGCTCTCAGGGTGGCCAGCCGCCCTTATGA 180  
 |||||  
 Db 121 CCATCGAGTACCTGGTTCGCTGACCGACGCGCTCTCAGGGTGGCCAGCCGCCCTTATGA 180  
 |||||

Qy 181 CTGTCCCGCGGGGTCGAGGTGCGGTGGAACCGACGACAT 223  
 |||||

Db 165 ACGTTCGCGGGGACCGAGGTCGCGGTGGAGACCGACGACAT 207  
 |||||

RESULT 14

AAS05205

ID AAS05205 standard; DNA; 208 BP.

XX AAS05205;

DT 07-SEP-2001 (first entry)

DE Mycobacterium tuberculosis rpoB gene fragment.

KW Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;  
 KW PCR-restriction fragment length polymorphism analysis; ds.

OS Mycobacterium tuberculosis.

PN WO200131061-A1.

PD 03-MAY-2001.

XX 27-OCT-2000; 2000WO-KR01223.

XX 27-OCT-1999; 99KR-0046795.

XX (ERUM-) ERUME BIOTECH CO LTD.

XX Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;

XX

WP1: 2001-300520/31.

New DNA fragments from the rpoB gene of mycobacteria, useful for diagnosis and identification of many mycobacterial species by restriction fragment length polymorphism.

Disclosure: Page 41; 50pp; English.

The present sequence for Mycobacterium tuberculosis rpoB gene fragment is 1 of 24 rpoB gene fragments (AA05201-AA0524) from various Mycobacterial species. These rpoB gene fragments can be used in the diagnosis and identification of Mycobacterium species using a novel PCR-restriction fragment length polymorphism analysis (PRA) method. The method comprises obtaining a restriction fragment length polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating, amplifying and digesting the DNA fragment from the microorganism to be identified and comparing the RFLP patterns from the known rpoB gene fragments with the unidentified fragment. The rpoB gene fragments are useful to identify a wide range of Mycobacterium species, e.g. for diagnosis or to obtain epidemiological and pathogenesis information for selection of appropriate therapies, including M. tuberculosis, M. leprae and non-tuberculous mycobacteria (NTM) encountered in subjects infected with human immunodeficiency virus (HIV). Analysis of the rpoB gene fragments is rapid, precise, simple and cost effective (only 1 PCR required), and can differentiate between many species in a single experiment, including those difficult to distinguish by usual biochemical tests. Also described are oligonucleotide probes (AA05227-AA05242) for detecting specific Mycobacterial species.

Sequence 208 BP; 46 A; 62 C; 67 G; 33 T; 0 other;

Query Match 55.5%; Score 123.8; DB 22; Length 208;  
 Best local Similarity 76.7%; Pred. No. 2.4e-20;  
 Matches 171; Conservative 0; Mismatches 37; Indels 15; Gaps 1

Qy 1 TCAAGGAAGCGCTACGACCTGGCCGCGTGGCCGCTACACGCTCAACGAAGCTGG 60  
 Db 1 TCAAGGAAGCGCTACGACCTGGCCGCGTGGCCGCTACACGCTCAACGAAGCTGG 60

Qy 61 GCCTGAACACCGATATCCGATCACACACACGCTGACCGAAGACGCTCGTCGCCA 120  
 Db 61 GCCTGATGTCGGCGAGCCCATCAGTCTCGACGCTGACCGAAGACGCTCGTCGCCA 120

Qy 121 CCATCGAGTACCTGGTTGCGCTGACACACGCTCTCAGGGTGGCCGCGCCCGTTATGA 180  
 Db 121 CCATCGAATATCTGGTCCGCTTCACGAGG-----GTCAGACCCACGATGA 165

Qy 181 CTCTCCCCCGCGGCTCGAGGTCCGCTGGAAACGACGACAT 223  
 Db 166 CCGTTCGGCGGCGCTCGAGGTCCGCTGGAAACGACGACAT 208

RESULT 15  
 AA05206  
 ID AA05206 standard; DNA; 208 BP.  
 AC AA05206:  
 XX  
 DT 07-SEP-2001 (first entry)  
 DE Mycobacterium terrae rpoB gene fragment.  
 XX Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;  
 KW PCR-restriction fragment length polymorphism analysis; ds.  
 XX Mycobacterium terrae.  
 OS  
 XX WO200131061-A1.  
 XX  
 PD 03-MAY-2001.  
 XX  
 XX 27-OCT-2000; 2000WO-KR01223.  
 XX

```

PR 27-OCT-1999;      99KR-0046795.
XX
XX (ERUM-) ERUME BIOTECH CO LTD.
XX PA
XX
XX Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ:
XX WPI: 2001-300520/31.
XX
XX New DNA fragments from the rpoB gene of mycobacteria, useful for
PT diagnosis and identification of many mycobacterial species by
PT restriction fragment length polymorphism -
XX
XX Claim 1: Page 42: 50pp: English.
XX
XX The present sequence for Mycobacterium terrae rpoB gene
CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
CC various Mycobacterial species. These rpoB gene fragments can be used
CC in the diagnosis and identification of Mycobacterium species using a
CC novel PCR-restriction fragment length polymorphism analysis (PRA)
CC method. The method comprises obtaining a restriction fragment length
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
CC amplifying and digesting the DNA fragment from the microorganism to
CC be identified and comparing the RFLP patterns from the known rpoB gene
CC fragments with the unidentified fragment. The rpoB gene fragments
CC are useful to identify a wide range of Mycobacterium species, e.g. for
CC diagnosis or to obtain epidemiological and pathogenesis information for
CC selection of appropriate therapies, including M. tuberculosis, M. leprae
CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene
CC fragments is rapid, precise, simple and cost effective (only 1 PCR
CC required), and can differentiate between many species in a single
CC experiment, including those difficult to distinguish by usual biochemical
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
CC detecting specific Mycobacterial species.
XX
XX Sequence 208 BP: 46 A; 62 C; 67 G; 33 T; 0 other:
SQ
Query Match          55.5%; Score 123.8; DB 22: Length 208;
Best local Similarity 76.7%; Pred No. 2.4e-20;
Matches 171; Conservative 0; Mismatches 37; Indels 15; Gaps 1
Qy 1 TCAGAGGAGAAGCGCTACGACTGGCCGCCGTTCGCGCGGTACAAGGTCAACGAAGAGTGG 60
Db 1 TCAGAGGAGAAGCGCTACGACTGGCCGCCGTTCGCGCGGTATTAAGGTCAACGAAGAGTGC 60
Qy 61 GCCTGAACACCGATCATCCGATCACCACACGACGCTGACCCGAACGACGTCGTCGCCCA 120
Db 61 GGCTGCATGTTCGGCGAGCCCATCACCTGTCACNCGCTGACCGAAGAACGCTGTGGCCA 120
Qy 121 CCATPCGAGTAGTACCTGGTTTCGCCTGCACACGCGCTCTCAAGGTGCCGCCGCCGCTTATGA 180
Db 121 CCATPCGAATAATCTGGTTCGCGTTGCAGAGG-----GTCAGACCCACGATGA 165
Qy 181 CTGTCCCCCGGGGGTCGAGGTGCCGCGTGGAACCGACGACAT 223
Db 166 CCGTTCGCGCGGCGTCGAGTGCCTCGGTGGAACCGACGACAT 208
Search completed: November 12, 2002, 16:50:12
Job time : 149.046 secs

```



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 12, 2002, 15:56:36 ; Search time 29.9052 Seconds  
(without alignments)  
2286.857 Million cell updates/sec

Title: US-09-697-123B-13

Perfect score: 223

Sequence: 1 tcaaggagaagcgctacgac.....ccggtggaaccgacgacat 223

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_NA.\*  
1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PCTUS\_COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	123.8	55.5	970	1	US-08-250-030-1
2	123.8	55.5	970	5	PCT-US95-06790-1
3	111	49.8	3447	2	US-08-313-185-57
4	111	49.8	3447	3	US-09-082-614A-57
5	41.2	18.5	3306	1	US-08-261-206A-71
6	38.8	17.4	491	1	US-08-133-711-36
7	37.2	16.7	3231	1	US-08-074-121-4
8	37.2	16.7	3231	5	PCT-US94-06447-4
9	36.6	16.4	734	4	US-09-221-017B-1070
10	36.6	16.4	765	3	US-08-718-904-79
11	36.2	16.2	1272	2	US-08-972-258-1
12	36.2	16.2	1272	4	US-09-263-128-1
13	36.2	16.2	2346	4	US-09-193-503B-4
14	36.2	16.2	2346	4	US-09-193-503B-5
15	36.2	16.2	2346	4	US-09-193-503B-8
16	36	16.1	1597	2	US-08-724-974A-1
17	35.4	15.9	1143	2	US-08-387-942C-19
18	35.4	15.9	12588	2	US-08-387-942C-1
19	35.4	15.9	4403765	4	US-09-103-840A-2
20	35	15.7	4848	4	US-08-955-957A-1
21	35	15.7	4848	4	US-08-955-957A-4
22	35	15.7	4848	4	US-08-955-957A-6
23	34.8	15.6	2064	1	US-08-343-428-1
24	34.6	15.5	1288	1	US-08-440-856A-9
25	34.4	15.4	1007	3	US-09-248-335-31
26	34.2	15.3	1251	4	US-09-105-537-38
27	34.2	15.3	5970	3	US-09-320-878-21

ALIGNMENTS

RESULT 1

US-08-250-030-1

; Sequence 1, Application US/08250030

; Patent No. 5643723

; GENERAL INFORMATION:

; APPLICANT: Persing, David H.

; TITLE OF INVENTION: Detection of a Genetic Locus Encoding

; TITLE OF INVENTION: Resistance to Rifampin in Mycobacterial Cultures and In

; NUMBER OF SEQUENCES: 15

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Schwegman, Lundberg & Woessner

; STREET: 3500 IDS Center

; CITY: Minneapolis

; STATE: MN

; COUNTRY: USA

; ZIP: 55402

; COMPUTER READABLE FORM:

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/250,030

; FILING DATE: 26-MAY-1994

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Muetling, Ann M.

; REGISTRATION NUMBER: 33,977

; REFERENCE/DOCKET NUMBER: 150.1050S1

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 612-339-0331

; TELEFAX: 612-339-3061

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 970 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA

US-08-250-030-1

Query Match 55.5% Score 123.8; DB 1; Length 970;

Best Local Similarity 76.7%; Pred. No. 1.4e-22;

Matches 171; Conservative 0; Mismatches 37; Indels 15; Gaps 1;

Oy 1 TCAAGGAGAAGCGCTACGACCTGGCCCGCTCGGCGCTACAAGGTCAACAAGAGCTGG 60

|||||

Db 26 TCAAGGAGAAGCGCTACGACCTGGCCCGCTCGGCTCGCTATTAAGTCAACAAGAGCTCG 85

|||||

61	GCCTGAACACCCAGTATCCGATCACCACACGACGCTGACCCGAAGAGACGCTGCTGC	120
Qy		
Db		
86	GGCTGCATGTCGGCGAGCCATCAGCTCTGTGACGCTGACCCGAAGAGACGCTGCTGC	145
121	TCATCAGTACCTGGTTTCGCGCTGCACACAGCGCTCTCAGGGTGCCAGGCCCGCTTATGA	180
Qy		
146	CCATCGAATATCTGCTCCGCTTTCACAGG-----GTCAGACACGATGA	190
Db		
181	CTCTCCCGCGCGGTGCGAGTGCCTGGAAACCCAGCACAT	223
Qy		
191	CCGTCTCCGGCGCGCTGAGGTCGCGTGGAAACCCAGCACAT	233
Db		

```

RESULT 2
PCT-US95-06790-1
: Sequence 1, Application PC/TUS9506790
: GENERAL INFORMATION:
: APPLICANT: Mayo Foundation for Medical Education and Research
: APPLICANT: and Hoffmann-La Roche Inc.
: TITLE OF INVENTION: Detection of a Genetic Locus Encoding
: TITLE OF INVENTION: Resistance to Rifampin
: NUMBER OF SEQUENCES: 15
: CORRESPONDENCE ADDRESS:
:

```

	Query Match	55.5%	Score 123-8;	DB 5;	Length 970;
	Best Local Similarity	76.7%	Pred. No. 1,4e-22;		
	Matches 171;	Conservative 0;	Mismatches 37;	Indels 15;	Gaps 1;
Qy	1	TCAAGGAACGCTACGACCTGC	CGCCGTCGCCGCTACAAGGTCAACAAGAAGCTGG	60	
Dd					
Dd	26	TCAAGGAACGCTACGACCTGC	CGCCGTCGGTATTAAGGTCAACAAGAAGCTCG	85	
Qy	61	GCCTGAACACCGCATCTCCGATCAC	CCACCAGCCTGCACCGAAGAGACGCTCGTCGCCA	120	
Dd					
Dd	86	GGCTGCATCTCGGGCAGGCCATC	ACGTCGTTCGACCTGCACCGAAGAGACGCTCGTGGCCA	145	
Qy	121	CCATCGACTACCTGGTTTCGCCCTGC	ACCACGCGCTCTCAGGGTGGCCAGCCCCCGCTATTGA	180	
Dd					
Dd	146	CCATCGAATATCTGGTTCGCTTGC	ACGAGG-----GTCAACACCACCATGA	190	
Qy	181	CTGTCCCCCGGGGTCGAGTGC	CGGTGGAAACCGACGACAT	223	
Dd	191	CCGTTCCGGGGCGGCTCCAGGTGC	CGGTGGAACCGACGACAT	233	

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

Query Match	49.8%	Score 111	DB 2	Length 3447
Best Local Similarity	73.1%	Pred. No. 2.5e-19		
Matches 163	Conservative 0	Mismatches 45	Indels 15	Gaps
Qy	1	TCAGGAGAAGCGCTACGACCTCGCCGCGTCCGCGCTACAAAGGTCAACAAGAACGTCGG	60	
Db	809	TCAPVGGAGAACGCTACGACCTCGCCAGGGTTCGTTACAGGTCACACAGACGTCG	868	
Qy	61	GCCTGAACACCGGATATCCCGATCACCACACGACGCTGACCGGAAGAGAGCTGTCGTCGCCA	120	
Db	869	GGTTGACCGCGGTGAGTTTCATCACGTCGTCACGCTGACCGAAGAGAGATGCTGTCGCCA	928	
Qy	121	CCATCGAGTACCTGGTTCGCGCTGCACACGCGCTCTCAGGGTGCGCCCGCGTTATCA	180	
Db	929	CCATAGAGTACCTGGTTCGTCGTCA-----TGAGGGTCAGTCGACACATCA	973	
Qy	181	CTGTCCCCCGGGGTGAGGTGCCGTGGAAACCGACGACAT	223	
Db	974	CTGTCCCAGGTGGGTGAGAGTGCCTGGAACCTCAGCATAT	1016	

RESULT 4  
US-09-082-614A-57

; Sequence 57, Application US/09082614A  
; Patent No. 6124098  
; GENERAL INFORMATION:  
; APPLICANT: Heym, Beate  
; APPLICANT: Cole, Stewart  
; APPLICANT: Young, Douglas  
; APPLICANT: Zhang, Ying  
; APPLICANT: Honore, Nadine  
; APPLICANT: Telenti, Amelio  
; APPLICANT: Hodmer, Thomas  
; TITLE OF INVENTION: Rapid Detection of Antibiotic Resistance  
; TITLE OF INVENTION: in Mycobacterium Tuberculosis  
; NUMBER OF SEQUENCES: 66  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &  
; ADDRESSEE: Dunner  
; STREET: 1300 I Street, N.W.  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20005-3315  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/082.614A  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/313,185  
; FILING DATE: 12-OCT-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Meyers, Kenneth J.  
; REGISTRATION NUMBER: 25,146  
; REFERENCE/DOCKET NUMBER: 02356.0068-00000  
; TELEPHONE: (202) 408-4000  
; TELEFAX: (202) 408-4400  
; INFORMATION FOR SEQ ID NO: 57:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3447 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
US-09-082-614A-57

Query Match 49.8%; Score 111; DB 3; Length 3447;  
Best Local Similarity 73.1%; Pred. No. 2.5e-19;  
Matches 163; Conservative 0; Mismatches 45; Indels 15; Gaps 1;  
QY 1 TCAAGGAGAGCGCTACGACCTGGCCGCGTACAGGTCACAAAGAGCTGG 60  
DB 809 TCAAGGAGAGCGCTACGACCTGGCCGCGTACAGGTCACAAAGAGCTGG 868  
QY 61 GCCTGAACACCGGATCATCCCATCACCACGACGCTGACCCGAAGAGAGCTGCGCCCA 120  
DB 869 GGTTCACGCGCGGTGAGTTATCAGCTGCTCCACGCTGACCCGAAGAGAGTGTGCGCCCA 928  
QY 121 CCATCGAGTACCTGGTTGCGCTGCACACGCTCTCAGGCGTGCAGGCCCGCTTATGA 180  
DB 929 CCATGAGTACCTGGTTGCTGCTCA-----TGAGGTCAGTTCACATGA 973  
QY 181 CTGTCCCGCGGCGGTGAGGTGCGCGTGGAAACCGACGACAT 223  
DB 974 CTGTCCCGCGGCGGTGAGGTGCGCGTGGAAACCGACGATAT 1016

RESULT 5  
US-08-261-206A-71  
; Sequence 71, Application US/08261206A

; Patent No. 5574007  
; GENERAL INFORMATION:  
; APPLICANT: Zushi, Mitichitaka  
; APPLICANT: Gomi, Komakazu  
; APPLICANT: Yamamoto, Shuji  
; APPLICANT: Suzuki, Koji  
; APPLICANT: Matsuda, Akio  
; TITLE OF INVENTION: A Polypeptide Capable of Interacting  
; TITLE OF INVENTION: with Thrombin  
; NUMBER OF SEQUENCES: 80  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Birch, Stewart, Kolasch & Birch  
; STREET: 301 N. Washington St.  
; CITY: Falls Church  
; STATE: Virginia  
; COUNTRY: USA  
; ZIP: 22046-0747  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/261.206A  
; FILING DATE:  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/740.492  
; FILING DATE: 03-AUG-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Svensson, Leonard R.  
; REGISTRATION NUMBER: 30330  
; REFERENCE/DOCKET NUMBER: 216-275P  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703-241-1300  
; TELEFAX: 703-241-2848  
; TELEX: 248345  
; INFORMATION FOR SEQ ID NO: 71:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3306 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; ORIGINAL SOURCE:  
; ORGANISM: Acremonium chrysogenum  
; FEATURE:  
; NAME/KEY: -  
; LOCATION: 1..3306  
; OTHER INFORMATION: /label= PGK gene  
; OTHER INFORMATION: /note= "Nucleotide sequence of region A in Figure  
; OTHER INFORMATION: 59. The sequence is presented as Figure 61."  
; FEATURE:  
; NAME/KEY: exon  
; LOCATION: 1252..1317  
; FEATURE:  
; NAME/KEY: exon  
; LOCATION: 1463..1883  
; FEATURE:  
; NAME/KEY: exon  
; LOCATION: 1948..2715  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: join(1252..1317, 1463..1883, 1948..2714)  
US-08-261-206A-71

Query Match 18.5%; Score 41.2; DB 1; Length 3306;  
Best Local Similarity 49.1%; Pred. No. 0.062;  
Matches 109; Conservative 0; Mismatches 113; Indels 0; Gaps 0;  
QY 1 TCAAGGAGAGCGCTACGACCTGGCCGCGTACAGGTCACAAAGAGCTGG 60  
DB 1455 TCACGCGAGGTGCGACTTCAACGTCGCCCTCGACGAGAAAGAGATCACCACCAACCGC 1514

Qy	61	GCCTGAACACCGCATCATCGATCACCACACAGACGCTCACCAGAAAGACGCTCTCGCCA	120
Db	1515	GCATCTGTCGTGCCCTCCCGACCATCACTACGGCTCGAGCATGGGCCAAGCCGCTCA	1574
Qy	121	CGATCCAGTACTGGTTCGCTTCGACACAGCCTCTCAGGTGGCCAGGCCGCCCTTATGA	180
Db	1575	TCTCATGTCCCACTTGGCCGCCCGCAAGCGGACCCCAACCCCAAGTACTCGCTGCAGC	1634
Qy	181	CTGTCCCGCGGGGTCGAGGTGCGCGTGGAAACCGACGACA	222
Db	1635	CCCTGCTCCCGCAGCTCGAGAAGCTGCTTCGGCAAGAGGGTCA	1676

## RESULT 6

US-08-133-711-36

: Sequence 36, Application US/0813711  
 : Patent No. 5523463  
 : GENERAL INFORMATION:  
 : APPLICANT: Zoig, Werner  
 : TITLE OF INVENTION: Methods and reagents for detection of  
 : TITLE OF INVENTION: pathogenic using superoxide dismutase gene  
 : TITLE OF INVENTION: Targeting  
 : NUMBER OF SEQUENCES: 44  
 : CORRESPONDENCE ADDRESS:  
 : ADDRESSEE: Hoffmann-La Roche Inc.  
 : STREET: 340 Kingsland Street  
 : CITY: Nutley  
 : STATE: NJ  
 : COUNTRY: U.S.A.

```
Query Match      17.4%; Score 38.8; DB 1; Length 491;
Best Local Similarity 48.6%; Pred. No. 0.19;
Matches 106; Conservative 0; Mismatches 112; Indels 0; Gaps 0;
```

Qy 61 GCCTGAACACCAGTTCATCCGATCACCACCACCGCTGCACCGAAGAAGAGTGCTGGCCA 120  
- - - - -  
Db 160 GTTGGGAAGAACCTTTTCCTCCCAACGGAGCGACAAGCCACCGCGCAACTCGCCGCGGCCCA 219

Oy	121	CCATCGAGTACCTGGTTCGGCTGCACACCCCTCTCAGGTGCCAGGCCCGCGTTATCA	180
Db	220	TCGACGAGGGGTTCGGGTCTCTCGACAAGTTCGTGCCAAATCCACGCCGCCGCCACCA	279
Oy	181	CTGTCCCGCCGGGGTTCGAGTCCGGCTGGAAACCGAC	218
Db	280	CGGTGCAGGGGTTCGGGTGGCGCGCGCGTGGGCTCGGAC	317

## RESULT 7

```

US-08-074-121-4
: Sequence 4, Application US/08074121
: Patent No. 5767362
: GENERAL INFORMATION:
: APPLICANT: Best, Elaine
: APPLICANT: Knauf, Vic C.
: TITLE OF INVENTION: Methods and Compositions for Modulating
: TITLE OF INVENTION: Lipid Content of Plant Tissues
: NUMBER OF SEQUENCES: 15
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Fish & Richardson
: STREET: 2200 Sand Hill Road, Suite 100
: CITY: Menlo Park
: STATE: California
: COUNTRY: US
: ZIP: 94025
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/074,121
: FILING DATE: 08-JUN-1993
: CLASSIFICATION: 800
: ATTORNEY/AGENT INFORMATION:
: NAME: Rae-Venter, Barbara
: REGISTRATION NUMBER: 32,750
: REFERENCE/DOCKET NUMBER: 05938/043001
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415) 854-5277
: TELEFAX: (415) 854-0875
: TELEX: 200154
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3231 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 861..1328
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 1349..2695
: US-08-074-121-4

```

ORIGINAL SOURCE: Mycobacterium kansasii/SOD gene  
US-08-133-711-36

```

: RESULT 8
: PCT-US94-06447-4
: Sequence 4, Application PC/TUS9406447
: GENERAL INFORMATION:
: APPLICANT: Calgene, Inc.
: TITLE OF INVENTION: Methods and Compositions for Modulating
: TITLE OF INVENTION: Lipid Content of Plant Tissues
: NUMBER OF SEQUENCES: 15
: CORRESPONDENCE ADDRESS:
: ADMP/SSFE: Weill, Golschal & Manges
: STREET: 2882 Sand Hill Road, Suite 280
: CITY: Menlo Park
: STATE: California
: COUNTRY: US
: ZIP: 94025
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: PCT/US94/06447
: FILING DATE: 06-JUN-1994
: CLASSIFICATION:
: ATTORNEY/AGENT INFORMATION:
: NAME: Rae-Venier, Barbara
: REGISTRATION NUMBER: 32,750
: REFERENCE/DOCKET NUMBER: CGNF-097/WO
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415) 926-6200
: TELEFAX: (415) 854-3713
: TELEX:
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3231 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 861..1328
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 1349..2695
: PCT-US94-06447-4

```

	Query Match	16.7%	Score 37.2:	DB 5:	length 3231:
	Best Local Similarity	51.2%:	Pred. No. 0.61:		
	Matches 87:	Conservative	0:	Mismatches 83:	Indels 0:
	Gaps	0:			
QY	29	CGTCGGCGCTACAAAGCTCAACAAGACCTGCGGCTCGAACACCGATCATTCGATCACCAC	88		
Db	1996	CGACGGCCAGGCAACGCCATCCACCTCGGCGACCGCGGACTGCTCCCTGCACGCGCGCCA	2055		
QY	89	CACGACGCTGACCCGAAGAAGACGTCGTCGCCACCATCGAGTACCTGGTTCGCCCTGCACCA	148		
Db	2056	CCAGAAGGTGATCGAAGAGGCGCCGCCCGCGGATCGACGAGAAGCTCGCCAGGAAGT	2115		
QY	149	GGCTCTCAGGTGGCCAGGCCCGCTTATGACTGTCCCGCGGGGTGCG	198		
Db	2116	CTTCGGCCGTGCGTCAGGCTCGATTCGAGATCGGCTACGCGCGCGCG	2165		

RESULT 9  
US-09-221-017B-1070  
: Sequence 1070, Application US/09221017B  
: Patent No. 6444799  
: GENERAL INFORMATION:  
: APPLICANT: ROSS, Bruce C.  
: TITLE OF INVENTION: P. GINGIVALIS NUCLEOTIDES AND USES THEREOF

1 NUMBER OF SEQUENCES: 1120  
 2  
 3 CORRESPONDENCE ADDRESS:  
 4 ADDRESSEE: MORRISON & FOERSTER  
 5 STREET: 755 PAGE MILL ROAD  
 6 CITY: Palo Alto  
 7 STATE: CA  
 8 COUNTRY: USA  
 9 ZIP: 94304-1018  
 10 COMPUTER READABLE FORM:  
 11 MEDIUM TYPE: Diskette  
 12 COMPUTER: IBM Compatible  
 13 OPERATING SYSTEM: Windows  
 14 SOFTWARE: FastSeq for Windows Version 2.0b  
 15 CURRENT APPLICATION DATA:  
 16 APPLICATION NUMBER: US/09/221,017B  
 17 FILING DATE: 23-DEC-1998  
 18 CLASSIFICATION:  
 19 PRIOR APPLICATION DATA:  
 20 APPLICATION NUMBER: PP1182  
 21 FILING DATE: 31-DEC-1997  
 22 PRIOR APPLICATION DATA:  
 23 APPLICATION NUMBER: PP1546  
 24 FILING DATE: 30-JAN-1998  
 25 PRIOR APPLICATION DATA:  
 26 APPLICATION NUMBER: PP2911  
 27 FILING DATE: 09-APR-1998  
 28 PRIOR APPLICATION DATA:  
 29 APPLICATION NUMBER: PCT/AU98/01023  
 30 FILING DATE: 10-DEC-1998  
 31 ATTORNEY/AGENT INFORMATION:  
 32 NAME: MONROY, Gladys H  
 33 REGISTRATION NUMBER: 32,430  
 34 REFERENCE/DOCKET NUMBER: 27340-20021.00  
 35 TELECOMMUNICATION INFORMATION:  
 36 TELEPHONE: 650-813-5600  
 37 TELEFAX: 650-494-0792  
 38 TELEX: 706141  
 39 INFORMATION FOR SEQ ID NO: 1070:  
 40 SEQUENCE CHARACTERISTICS:  
 41 LENGTH: 734 base pairs  
 42 TYPE: nucleic acid  
 43 STRANDEDNESS: double  
 44 TOPOLOGY: circular  
 45 MOLECULE TYPE: DNA (genomic)  
 46 HYPOTHETICAL: NO  
 47 ANTI-SENSE: UNKNOWN  
 48 ORIGINAL SOURCE:  
 49 ORGANISM: PORYPHYROMONAS GINGIVALIS  
 50 FEATURE:  
 51 NAME/KEY: misc\_feature  
 52 LOCATION: 1..734  
 53 US-09-221-017B-1070

[illegible]

RESULT 10  
US-08-718-904-79  
; Sequence 79, Application US/08718904

```

: Patent No. 6037329
:
: GENERAL INFORMATION:
:
: APPLICANT: Baird, J. Andrew
: APPLICANT: Chandler, Lois Ann
: APPLICANT: Sosnowski, Barbara A.
: TITLE OF INVENTION: COMPOSITIONS CONTAINING NUCLEIC ACIDS AND LIGANDS FOR THERAPE
:
: NUMBER OF SEQUENCES: 128
:
: CORRESPONDENCE ADDRESS:
:
: ADDRESSEE: SEED AND BERRY LLP
: STREET: 6300 Columbia Center, 701 Fifth Avenue
: CITY: Seattle
: STATE: Washington
: COUNTRY: USA
: ZIP: 98104-7092
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
:
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
:
: CURRENT APPLICATION DATA:
:
: APPLICATION NUMBER: US/08/718,904
: FILING DATE: 24-SEP-1996
: CLASSIFICATION: 424
:
: ATTORNEY/AGENT INFORMATION:
: NAME: No. 6037329tenburg Ph.D., Carol
: REGISTRATION NUMBER: 39,317
: REFERENCE/DOCKET NUMBER: 760100.415C1
:
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (206) 622-4900
: TELEFAX: (206) 682-6031
:
: INFORMATION FOR SEQ ID NO: 79:
:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 765 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
:
: FEATURE:
:
: NAME/KEY: CDS
: LOCATION: 1..762
:
: OTHER INFORMATION: /product= "Mammalian codon
: OTHER INFORMATION: optimized saporin"
:
: US-08-718-904-79

```

	Query Match	16.4%	Score 36.6;	DB 3;	Length 765;
	Best Local Similarity	53.1%;	Pred. No.	0.71;	
Matches	78:	Conservative	0;	Mismatches	69; Indels 0; Gaps 0;
Qy	6	GAGAAAGCGCTACGACCTGGCCCGCGTCGGCGCTTACRAGGTCAACAAGAAGCTGGGCCTG	65		
Db	208	GACAACTGTACTGGTGCGCTACTCTGCCATGACACAACACGTGAACCGCGCCTAC	267		
Qy	56	AACACCAGTATCCGATCACCAACACACGCTGACCCGAAGAAGAGCTGTCGCCACCATC	125		
Db	268	TACTTCAAGTCGAGATCACTCCGCCGAGGTGACCCGCTTTCCCTGAGGCCACCA	327		
Qy	126	GAGTACCTGGTTGCGCTGCAACAGCC	152		
Db	328	GCCAACCAAGAGGCCCTGGAGTACACC	354		

```

RESULT 11
US-08-972-258-1
: Sequence 1, Application US/08972258
: Patent No. 5929301
: GENERAL INFORMATION:
: APPLICANT: Baszczynski, Chris
: APPLICANT: Bowen, Benjamin A.
: APPLICANT: Drummond, Bruce J.
: APPLICANT: Gordon-Kamm, William
: APPLICANT: Peterson, David J.
: APPLICANT: Sandahi, Gary A.
: APPLICANT: Tagliani, Laura A.
: APPLICANT: Zhao, Zuo-Yu

```

```

: TITLE OF INVENTION: No. 5929301el Nucleic Acid Sequence Encoding FLP
: Patent No. 5929301
: TITLE OF INVENTION: Recombinase and Method of Using Same
: NUMBER OF SEQUENCES: 4
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: W. Murray Spruill
: STREET: 3605 Glenwood Ave. Suite 310
: CITY: Raleigh
: STATE: NC
: COUNTRY: US
: ZIP: 27622
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/972,258
: FILING DATE:
: CLASSIFICATION: 800
: ATTORNEY/AGENT INFORMATION:
: NAME: Spruill, W. Murray
: REGISTRATION NUMBER: 32,943
: REFERENCE/DOCKET NUMBER: 5718-28
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 919 420 2202
: TELEFAX: 919 881 3175
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1272 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: ORIGINAL SOURCE:
: ORGANISM: Synthetic sequence (optimized)
: JS-08-972-258-1

```

Query Match	16.2%	Score 36.2	DB 2:	Length 1272;
Best Local Similarity	48.8%	Pred. No. 0.96;		
Matches 98:	Conservative 0;	Mismatches 103;	Indels 0;	Gaps 0;
Qy 20	CCTGCCCGCTCGGCCGCTACAGGTCACCAAGAAGCTGGCCCTGAACACCGCATCACC	79		
Db	540	CCTCTTCTCGCCACCTTCATCAACTCGGCAGGTTCTCACACATCAGAACGTTGACCC	599	
Qy 80	GATCACCACACGAGCGTGCAGGAAGAGAGCTGCTGCCACCATCTAGTACCTGGTTCCG	139		
Db	600	CAAGTCCTTCAAGCTCGTGCAGAACAGTACCTCGGCGTGATCATCCAGTGCCTCGTGAC	659	
Qy 140	CCTGCACCACCGCTTCTCAGGTGGCCAGGCCCCCGTTATGACTGTCGCCCGCGGGTGCA	199		
Db	660	CGAGACCAAGACCTCCGTGTCAGGCACATCTACTTCTTCGCGTCGGCGCAGGATCGA	719	
Qy 200	GGTCCGGTGGAAACCGACGA	220		
Db 720	CCCCCTCGTGACCTCGACGA	740		

```

RESULT 12
US-09-263-128-1
: Sequence 1, Application US/09263128
: Patent No. 6175058
: GENERAL INFORMATION:
: APPLICANT: Baszczyński, Chris
: APPLICANT: Bowen, Benjamin A.
: APPLICANT: Drummond, Bruce J.
: APPLICANT: Gordon-Kamm, William J
: APPLICANT: Peterson, David J.
: APPLICANT: Sandahl, Gary A.
: APPLICANT: Tagliani, Laura A.
: APPLICANT: Zhao, Zuo-Yu
: TITLE OF INVENTION: No. 6175058rel

```



OTHER INFORMATION: Description of Artificial Sequence: sequence  
OTHER INFORMATION: encoding moCre:FLPm, Cre from Bacteriophage P1 and  
OTHER INFORMATION: FLP from Saccharomyces, both maize preferred  
OTHER INFORMATION: codons  
NAME/KEY: CDS  
LOCATION: (1)..(2346)  
US-09-193-503B-5

Query Match 16.2%; Score 36.2; DB 4; Length 2346;  
Best Local Similarity 48.8%; Prod. No. 1;  
Matches 98; Conservative 0; Mismatches 103; Indels 0; Gaps 0;  
QY 20 CCTGCCCGCGGTGCGCGCTACAAAGTCAACAAAGAGCTGGCGCTGAACACCGATCATCC 79  
DB 1614 CCCTCTCCCTCGCCACCTTCATCAACTCCGCGCAGGTTCTCAGACATCAAGAACCTGGACCC 1673  
QY 80 GATCACCACACGACGCTGACCAAGAGAGAGCTGCTGCGCCACCATCGAGTACCTGGTTGC 139  
DB 1674 CAAGTCCCTTCAAGCTGCTGCAAGAACAGTAGTACCTCGCGGTGATCATCGAGTGCCTCGTGAC 1733  
QY 140 CCTGCACACGCCCTCTCAGGCTGCGCAGGCCCGCGTTTATGACTGTCCCGCGCGGGGTGCA 199  
DB 1734 CGACCAAGACCTCCGTGTCAGGCACATCTACTTCTTCTCCGCTCGCGCGAGGATCGA 1793  
QY 200 GGTCCCGGTGGAACCGACGA 220  
DB 1794 CCCCCTCGTGATCGACGA 1814

RESULT 15

US-09-193-503B-8  
Sequence 8, Application US/09193503B  
Patent No. 6262341  
GENERAL INFORMATION:  
APPLICANT: Baszczynski, Christopher L.  
APPLICANT: Lyznik, Leszek A.  
APPLICANT: Gordon-Kamm, William J.  
APPLICANT: Guan, Xueni  
APPLICANT: Rao, Guru  
APPLICANT: Tagliani, Laura A.  
TITLE OF INVENTION: A No. 6262341el Method For The Integration Of Foreign DNA Into  
TITLE OF INVENTION: Eukaryotic Genomes  
FILE REFERENCE: 5718-66 (amended listing)  
CURRENT APPLICATION NUMBER: US/09/193,503B  
PRIOR FILING DATE: 1998-11-17  
CURRENT APPLICATION NUMBER: 60/099,435  
PRIOR FILING DATE: 1998-09-08  
PRIOR APPLICATION NUMBER: 60/056,627  
PRIOR FILING DATE: 1997-11-18  
PRIOR APPLICATION NUMBER: 60/065,613  
PRIOR FILING DATE: 1997-11-18  
NUMBER OF SEQ ID NOS: 11  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 8  
LENGTH: 2346  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Sequence  
OTHER INFORMATION: encoding a FLPm:Cre polypeptide, FLP from  
OTHER INFORMATION: Saccharomyces (maize preferred codons), and Cre  
NAME/KEY: CDS  
LOCATION: (1)..(2346)  
US-09-193-503B-8

Query Match 16.2%; Score 36.2; DB 4; Length 2346;  
Best Local Similarity 48.8%; Prod. No. 1;  
Matches 98; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

QY 20 CCTGCCCGCGGTGCGCGCTACAAAGTCAACAAAGAGCTGGCGCTGAACACCGATCATCC 79  
DB 540 CCTCTCTCCGCCACCTTCATCAACTCCGCGCAGGTTCTCAGACATCAAGAACCTGGACCC 599

QY 80 GATCACCACGACGCTGACCCGAAGACAGACGCTGCGCCACCATCGAGTACCTGGTTGC 139  
DB 600 CAAGTCTTCAAGCTCGTGCAGAACAAAGTACCTCGCGGTGATCATCCAGTGCCTCGTGCAC 659  
QY 140 CCTGCACACGCGCTCTCAGGGTGGCGCAGGCCCGCGTTTATGACTGTCCCGCGCGGGGTGCA 199  
DB 660 CGAGACCAAGAGCTCGGTGTCCAGGCACATCTACTTCTTCTCCGCTCGCGCGGAGGATCGA 719  
QY 200 GGTCCCGGTGGAACCGACGA 220  
DB 720 CCCCCTCGTGATCGACGA 740

Search completed: November 12, 2002, 20:14:30  
Job time : 51.9052 secs







GENERAL INFORMATION:  
APPLICANT: MOECKEL, Bettina  
APPLICANT: BATHE, Brigitte  
APPLICANT: HERMANN, Thomas  
APPLICANT: PFEFFERLE, Walter  
APPLICANT: KINDER, Michael  
TITLE OF INVENTION: NUCLEOTIDE SEQUENCES WHICH CODE FOR THE rpoB GENE:  
FILE REFERENCE: 204212USOX  
CURRENT APPLICATION NUMBER: US/09/887.052  
CURRENT FILING DATE: 2001-06-25  
PRIOR APPLICATION NUMBER: DE10107229.5  
PRIOR FILING DATE: 2001-02-16  
NUMBER OF SEQ ID NOS: 8  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 5  
LENGTH: 5099  
TYPE: DNA  
ORGANISM: Corynebacterium glutamicum  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (702)..(4196)  
US-09-887-052-5

Query Match 28.2%; Score 62.8; DB 10; Length 5099;  
Best Local Similarity 71.0%; Pred. No. 6.1e-08;  
Matches 98; Conservative 0; Mismatches 37; Indels 3; Gaps 1;

QY 9 AAGCGCTACGACCTGGCGCGCTGCGCTACAAAGGTCAACAAAGAGCTGGCGCTGAAC 68  
|||||  
DB 1578 AAGCGCTACGACCTGGCTGGCTTACAAAGATCAACCGCAAGCTCGCGCTGGT 1637  
|||||

QY 64 ACCGATATCCGATCACCACACGACGCTGACCGAAGAGAGCTGTCGCCACCATCGAG 128  
|||||

DB 1638 GCGACACGAGTGGTTTCAT---GACTCTTACTGAAGAGGACATCGCAACCAACCATCGAG 1694  
|||||

QY 129 TACCTGGTTCGCTGCAC 146  
|||||

DB 1695 TACCTGGTTCGCTGCAC 1712  
|||||

RESULT 6  
US-10-062-254-109  
Sequence 109, Application US/10062254  
Patent No. US20020138882A1  
GENERAL INFORMATION:  
APPLICANT: Cahoon, Edgar B  
APPLICANT: Cahoon, Rebecca E  
APPLICANT: Falco, Saverio Carl  
APPLICANT: Fang, Yiwen  
APPLICANT: Hantke, Sabine S.  
APPLICANT: Lee, Jian-Ming  
APPLICANT: Li, Zhongsen  
APPLICANT: Miao, Guo-Hua  
APPLICANT: Morgante, Michele  
APPLICANT: Niu, Xiping  
APPLICANT: Odell, Joan  
APPLICANT: Rafalski, Antoni  
APPLICANT: Sakai, Hajime  
APPLICANT: Zheng, Peizhong  
APPLICANT: Zhu, Qun  
TITLE OF INVENTION: Polynucleotides Encoding Proteins Involved In Plant Metabolism  
FILE REFERENCE:  
CURRENT APPLICATION NUMBER: US/10/062,254  
CURRENT FILING DATE: 2002-02-01  
PRIOR APPLICATION NUMBER: 09/630,346  
PRIOR FILING DATE: 2000-07-28  
PRIOR APPLICATION NUMBER: 60/146511  
PRIOR FILING DATE: 1999-07-30  
PRIOR APPLICATION NUMBER: 60/156006  
PRIOR FILING DATE: 1999-09-23  
PRIOR APPLICATION NUMBER: 60/156899  
PRIOR FILING DATE: 1999-09-30  
PRIOR APPLICATION NUMBER: 60/157287

PRIOR FILING DATE: 1999-10-01  
PRIOR APPLICATION NUMBER: 60/169767  
PRIOR FILING DATE: 1999-12-09  
PRIOR APPLICATION NUMBER: 60/171054  
PRIOR FILING DATE: 1999-12-16  
PRIOR APPLICATION NUMBER: 60/172958  
PRIOR FILING DATE: 1999-12-21  
PRIOR APPLICATION NUMBER: 60/171515  
PRIOR FILING DATE: 1999-12-22  
PRIOR APPLICATION NUMBER: 60/173535  
PRIOR FILING DATE: 1999-12-29  
NUMBER OF SEQ ID NOS: 375  
SOFTWARE: Microsoft Office 97  
SEQ ID NO 109  
LENGTH: 704  
TYPE: DNA  
ORGANISM: Zea mays  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (640)  
NAME/KEY: unsure  
LOCATION: (663)  
NAME/KEY: unsure  
LOCATION: (688)  
NAME/KEY: unsure  
LOCATION: (692)  
NAME/KEY: unsure  
LOCATION: (696)  
US-10-062-254-109

Query Match 19.0%; Score 42.4; DB 12; Length 704;  
Best Local Similarity 58.9%; Pred. No. 0.0094;  
Matches 73; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 4 AGGAGAAGCGCTAGCAGCTGGCGCGCTGCGCGCTCAAGGTCAACAAGAGCTGGGCC 63  
|||  
DB 275 AGCTGATGCGCGCAGCACCTGCCACCGATGCCGCGAAGAGGTCCCAAGAGGTGGGCC 334  
|||

QY 64 TGAACACCGATCATCCGATCACCACGACGCTGACCGAAGAGAGCTGTCGCCACCA 123  
|||  
DB 335 TCGTCTACCGCTCAACATCGCGCCCAAGAGATCGCGGTTGATGAGGAGATCTTCGTGG 394  
|||

QY 124 TCGA 127  
|||||

DB 395 TCGA 398  
|||||

RESULT 7  
US-10-062-254-111  
Sequence 111, Application US/10062254  
Patent No. US20020138882A1  
GENERAL INFORMATION:  
APPLICANT: Cahoon, Edgar B  
APPLICANT: Cahoon, Rebecca E  
APPLICANT: Falco, Saverio Carl  
APPLICANT: Fang, Yiwen  
APPLICANT: Hantke, Sabine S.  
APPLICANT: Lee, Jian-Ming  
APPLICANT: Li, Zhongsen  
APPLICANT: Miao, Guo-Hua  
APPLICANT: Morgante, Michele  
APPLICANT: Niu, Xiping  
APPLICANT: Odell, Joan  
APPLICANT: Rafalski, Antoni  
APPLICANT: Sakai, Hajime  
APPLICANT: Zheng, Peizhong  
APPLICANT: Zhu, Qun  
TITLE OF INVENTION: Polynucleotides Encoding Proteins Involved In Plant Metabolism  
FILE REFERENCE:  
CURRENT APPLICATION NUMBER: US/10/062,254  
CURRENT FILING DATE: 2002-02-01  
PRIOR APPLICATION NUMBER: 09/630,346  
PRIOR FILING DATE: 2000-07-28







```
: LENGTH: 642
: TYPE: DNA
: ORGANISM: Trilicium aestivum
: FEATURE:
: NAME/KEY: unsure
: LOCATION: (34)
: NAME/KEY: unsure
: LOCATION: (624)
US-10-062-254-123

Query Match
  Best Local Similarity 16.1%; Score 36; DB 12; Length 642;
  Matches 69; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

Qy 4 AGGAGAAGCGGTACGACCTGGCCCGGCTGCGCCGCTACAAAGGTCAACAAGAGCTGGGCC 63
Db 236 AGCAGATGAGCGCCACCTCGCCACCGACCGCGCGGCGGCAAGGAGATCGCCAAAGGTCGGCT 295

Qy 64 TGAACACCGATATCCCGATCACCGATCACCGACGCTGACCGAAGAAGAGCTGTCGCCACCA 123
Db 296 TCGTCTACCAAGCTCAACATCTCGCCCAAGAAGATGGGCGTGAGGAGATCTTCGTGCG 355

Qy 124 TCGA 127
Db 356 TCGA 359

RESULT 15
US-09-815-242-7705
: Sequence 7705, Application US/09815242
: Patent No. US20020061569A1
: GENERAL INFORMATION:
: APPLICANT: Haselbeck, Robert
: APPLICANT: Ohlsen, Kari L.
: APPLICANT: Zyskind, Judith W.
: APPLICANT: Wall, Daniel
: APPLICANT: Trawick, John D.
: APPLICANT: Carr, Grant J.
: APPLICANT: Yamamoto, Robert T.
: APPLICANT: Xu, H. Howard
: TITLE OF INVENTION: Identification of Essential Genes in
: FILE OF INVENTION: Prokaryotes
: FILE REFERENCE: ELITRA 011A
: CURRENT APPLICATION NUMBER: US/09/815,242
: PRIOR APPLICATION NUMBER: 60/191,078
: PRIOR FILING DATE: 2000-03-21
: PRIOR APPLICATION NUMBER: 60/206,848
: PRIOR FILING DATE: 2000-05-23
: PRIOR APPLICATION NUMBER: 60/207,727
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: 60/242,578
: PRIOR FILING DATE: 2000-10-23
: PRIOR APPLICATION NUMBER: 60/253,625
: PRIOR FILING DATE: 2000-11-27
: PRIOR APPLICATION NUMBER: 60/257,931
: PRIOR FILING DATE: 2000-12-22
: PRIOR APPLICATION NUMBER: 60/269,308
: PRIOR FILING DATE: 2001-02-16
: NUMBER OF SEQ ID NOS: 14110
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 7705
: LENGTH: 918
: TYPE: DNA
: ORGANISM: Pseudomonas aeruginosa
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (1)...(918)
US-09-815-242-7705

Query Match
  Best Local Similarity 16.1%; Score 36; DB 10; Length 918;
  Matches 105; Conservative 0; Mismatches 115; Indels 0; Gaps 0;
```

```
Qy 2 CAAGGAGAAGCGGTACGACCTGGCCCGGCTGCGCCGCTACAAAGGTCAACAAGAGCTGGG 61
Db 180 CGAGGTGCAGGCGGTCTACGTTCGATACGCCCGGCTGCAACAAGAGCGCGAGAAGGCCCT 239
Qy 62 CCTGAACACCGATCATCCGATCACCACCGACGCTGACCGAAGAAGAGCTGCTGCCAC 121
Db 240 CAACCGCTACATGAACCGTACCGCCAGCGGCGCATTTGAAGGATGTCGACGTTGATCTT 299
Qy 122 CATCGAGTACCTGGTTTCGCTGCACACGCTCTCTCAGGTTGGCCAGGCCGCCGTTATGAC 181
Db 300 CGTGTGCGACCGCACCCGCTGGACCGAAGAGGACCATGTCCTTGAGCGCGTGCAGTA 359
Qy 182 TGTCCCCGCGGGTTCGAGGTGCCGCTGGAACCGCAC 221
Db 360 CGTCAGTTGCCCGGTGCTGATCGCGGTGAACAAGACCGAC 399
```

Search completed: November 12, 2002, 16:58:16  
Job time : 33.6867 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.  
OM nucleic - nucleic search, using sw model  
Run on: November 12, 2002, 15:44:35 : Search time 136.222 Seconds  
(without alignments)  
3438.621 Million cell updates/sec

Title: US-09-697-123B-8  
Perfect score: 208  
Sequence: 1 tcaagagaagcgctacgac.....ccggtgaaaccgacgacat 208

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_101002:\*  
1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:\*  
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:\*  
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:\*  
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:\*  
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:\*  
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:\*  
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:\*  
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:\*  
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:\*  
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:\*  
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:\*  
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:\*  
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:\*  
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:\*  
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:\*  
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:\*  
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:\*  
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:\*  
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:\*  
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:\*  
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:\*  
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:\*  
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:\*  
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES				
Result No.	Score	Query Match	Length DB ID	Description
1	208	100.0	208 22 AAS05208	Mycobacterium kans
2	174.4	83.8	208 22 AAS05217	Mycobacterium celi
3	171.2	82.3	208 22 AAS05215	Mycobacterium aviu
4	169.6	81.5	208 22 AAS05203	Mycobacterium gord
5	169.6	81.5	208 22 AAS05211	Mycobacterium mari
6	168	80.8	208 22 AAS05201	Mycobacterium gord
7	168	80.8	208 22 AAS05219	Mycobacterium fort
8	166.4	80.0	208 22 AAS05205	Mycobacterium tube
9	166.4	80.0	208 22 AAS05206	Mycobacterium terr

10	166.4	80.0	208 22 AAS05210	Mycobacterium ulce
11	166.4	80.0	208 22 AAS05216	Mycobacterium bovi
12	166.4	80.0	970 17 AAT09676	Mycobacterium tube
13	166.4	80.0	3519 22 AAH51976	Mycobacterium tube
14	166.4	80.0	3534 22 AAH02079	Mycobacterium tube
15	166.4	80.0	3853 21 AAA74651	M. tuberculosis rp
16	166.4	80.0	3853 21 AAA89994	M. tuberculosis rp
17	164.8	79.2	208 22 AAS05222	Mycobacterium afri
18	161.6	77.7	208 22 AAS05224	Mycobacterium xeno
19	160	76.9	208 22 AAS05202	Mycobacterium gord
20	159.2	76.5	207 22 AAS05204	Mycobacterium gord
21	158.4	76.2	208 22 AAS05223	Mycobacterium haem
22	152.6	73.4	223 22 AAS05213	Mycobacterium gast
23	151.2	72.7	207 22 AAS05212	Mycobacterium szul
24	150.4	72.3	205 22 AAS05220	Mycobacterium intr
25	142.4	68.5	214 22 AAS05221	Mycobacterium absc
26	142.4	68.5	3447 14 AAQ51532	M.leprae rpoB gene
27	139.2	66.9	214 22 AAS05218	Mycobacterium malm
28	134.2	64.5	211 22 AAS05214	Mycobacterium flav
29	132.8	63.8	214 22 AAS05207	Mycobacterium chel
30	130.2	62.6	223 22 AAS05209	Mycobacterium scro
31	80	38.5	3495 22 AAH65512	C glutamicum codin
32	80	38.5	349980 22 AAH68525	C glutamicum codin
33	72.4	34.8	27426 23 AAS59541	Propionibacterium
34	43.2	20.8	4956 23 ABL09357	Drosophila melanog
35	43.2	20.8	35633 23 ABL09356	S. tendae nikkomyc
36	40.6	19.5	2918 21 AA243922	Streptomyces tende
37	40.6	19.5	2919 21 AA244486	Recombinant equirr
38	40.2	19.3	1242 18 AAT78176	M. capsulatus gene
39	40	19.2	2826 24 ABQ90502	M. capsulatus gene
40	40	19.2	49999 24 ABQ90984	50K-cellulase gene
41	39.4	18.9	1894 18 AAT66542	Maize glutathione-
42	39	18.8	1100 21 AAA76277	DNA sequence encod
43	39	18.8	1167 10 AAQ91083	Maize nitrite redu
44	39	18.8	1844 12 AAQ12000	P. putida KT2440-a
45	38.6	18.6	2142 22 AAF61045	

ALIGNMENTS

RESULT 1  
AAS05208  
ID AAS05208 standard; DNA; 208 BP.

XX AAS05208;

AC AAS05208;

DT 07-SEP-2001 (first entry)

XX Mycobacterium kansasii rpoB gene fragment.

DE Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;

XX PCR-restriction fragment length polymorphism analysis; ds.

KW Mycobacterium kansasii.

OS Mycobacterium kansasii.

PN WO200131061-A1.

XX 03-MAY-2001.

PD 27-OCT-2000; 2000WO-KR01223.

XX 27-OCT-1999; 99KR-0046795.

PR (ERUM-) ERUME BIOTECH CO LTD.

PA Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;

XX WPI; 2001-300520/31.

XX New DNA fragments from the rpoB gene of mycobacteria, useful for

PT diagnosis and identification of many mycobacterial species by

PT restriction fragment length polymorphism -



```

XX PS Claim 1; Page 42; 50pp; English.
XX CC The present sequence for Mycobacterium kansasii rpoB gene
CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
CC various Mycobacterial species. These rpoB gene fragments can be used
CC in the diagnosis and identification of Mycobacterium species using a
CC novel PCR-restriction fragment length polymorphism analysis (PRA)
CC method. The method comprises obtaining a restriction fragment length
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
CC amplifying and digesting the DNA fragment from the microorganism to
CC be identified and comparing the RFLP patterns from the known rpoB gene
CC fragments with the unidentified fragment. The rpoB gene fragments
CC are useful to identify a wide range of Mycobacterium species, e.g. for
CC diagnosis or to obtain epidemiological and pathogenesis information for
CC selection of appropriate therapies, including M. tuberculosis, M. leprae
CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene
CC fragments is rapid, precise, simple and cost effective (only 1 PCR
CC required), and can differentiate between many species in a single
CC experiment, including those difficult to distinguish by usual biochemical
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
CC detecting specific Mycobacterial species.
XX SQ Sequence 208 BP; 51 A; 65 C; 65 G; 27 T; 0 other;
      Query Match          100.0%; Score 208; DB 22; Length 208;
      Best Local Similarity 100.0%; Pred. No. 6.6e-41;
      Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TCAAGGAGAAGCGCTACGACCTGCGCCGCTGTCGGCCGATACAAAGTCAACAAGACTGG 60
DB 1 TCAAGGAGAAGCGCTACGACCTGCGCCGCTGTCGGCCGATACAAAGTCAACAAGACTGG 60
QY 61 GCCTGAACACCAATCATCCGATCACCACGACGACGCTGACCGAAGAGACGCTGTCGCCA 120
DB 61 GCCTGAACACCAATCATCCGATCACCACGACGACGCTGACCGAAGAGACGCTGTCGCCA 120
QY 121 CCATCGAGTATCTGGTCCGCTGTCAGAGGCGCCAGCCAGCATGACCGTCCGGGGCGGG 180
DB 121 CCATCGAGTATCTGGTCCGCTGTCAGAGGCGCCAGCCAGCATGACCGTCCGGGGCGGG 180
QY 181 TCGAGGTGCCGGTGGAAACCGACGACAT 208
DB 181 TCGAGGTGCCGGTGGAAACCGACGACAT 208
RESULT 2
AAS05217
ID AAS05217 standard; DNA; 208 BP.
AC AAS05217;
XX 07-SEP-2001 (first entry)
XX Mycobacterium celatum rpoB gene fragment.
XX Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;
XX PCR-restriction fragment length polymorphism analysis; ds.
XX Mycobacterium celatum.
OS WO200131061-A1.
XX 03-MAY-2001.
XX 27-OCT-2000; 2000WO-KR01223.
XX 27-OCT-1999; 99KR-0046795.
XX (ERUM-) ERUME BIOTECH CO LTD.
XX Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;
PI 27-OCT-2000; 2000WO-KR01223.

```

```

XX WPI; 2001-300520/31.
XX New DNA fragments from the rpoB gene of mycobacteria, useful for
XX diagnosis and identification of many mycobacterial species by
XX restriction fragment length polymorphism -
PS Claim 1; Page 45; 50pp; English.
XX The present sequence for Mycobacterium celatum rpoB gene
XX fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
XX various Mycobacterial species. These rpoB gene fragments can be used
XX in the diagnosis and identification of Mycobacterium species using a
XX novel PCR-restriction fragment length polymorphism analysis (PRA)
XX method. The method comprises obtaining a restriction fragment length
XX polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
XX amplifying and digesting the DNA fragment from the microorganism to
XX be identified and comparing the RFLP patterns from the known rpoB gene
XX fragments with the unidentified fragment. The rpoB gene fragments
XX are useful to identify a wide range of Mycobacterium species, e.g. for
XX diagnosis or to obtain epidemiological and pathogenesis information for
XX selection of appropriate therapies, including M. tuberculosis, M. leprae
XX and non-tuberculous mycobacteria (NTM) encountered in subjects infected
XX with human immunodeficiency virus (HIV). Analysis of the rpoB gene
XX fragments is rapid, precise, simple and cost effective (only 1 PCR
XX required), and can differentiate between many species in a single
XX experiment, including those difficult to distinguish by usual biochemical
XX tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
XX detecting specific Mycobacterial species.
XX SQ Sequence 208 BP; 48 A; 71 C; 64 G; 25 T; 0 other;
      Query Match          83.8%; Score 174.4; DB 22; Length 208;
      Best Local Similarity 89.9%; Pred. No. 7.2e-33;
      Matches 187; Conservative 0; Mismatches 21; Indels 0; Gaps 0;
QY 1 TCAAGGAGAAGCGCTACGACCTGCGCCGCTGTCGGCCGATACAAAGTCAACAAGAGCTGG 60
DB 1 TCAAGGAGAAGCGCTACGACCTGCGCCGCTGTCGGCCGATACAAAGTCAACAAGAGCTGG 60
QY 61 GCCTGAACACCAATCATCCGATCACCACGACGACGCTGACCGAAGAGACGCTGTCGCCA 120
DB 61 GCCTGAACACCGCTGCCCGATCAGGACGACCTCTCACCAGAGAGGACGCTGTCGCCA 120
QY 121 CCATCGAGTATCTGGTCCGCTGTCAGAGGCGCCAGCCAGCATGACCGTCCGGGGCGGG 180
DB 121 CCATCGAGTATCTGGTCCGCTGTCAGAGGCGCCAGCCAGCATGACCGTCCGGGGCGGG 180
QY 181 TCGAGGTGCCGGTGGAAACCGACGACAT 208
DB 181 TCGAGGTGCCGGTGGAAACCGACGACAT 208
RESULT 3
AAS05215
ID AAS05215 standard; DNA; 208 BP.
XX AAS05215;
XX 07-SEP-2001 (first entry)
XX Mycobacterium avium rpoB gene fragment.
XX Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;
XX PCR-restriction fragment length polymorphism analysis; ds.
XX Mycobacterium avium.
OS WO200131061-A1.
XX 03-MAY-2001.
XX 27-OCT-2000; 2000WO-KR01223.

```



DE Mycobacterium marinum rpoB gene fragment.  
 KW Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;  
 KW PCR-restriction fragment length polymorphism analysis; ds.  
 XX Mycobacterium marinum.  
 OS  
 XX  
 PN WO200131061-A1.  
 XX  
 XX 03-MAY-2001.  
 XX  
 XX 27-OCT-2000; 2000WO-KR01223.  
 PF  
 XX 27-OCT-1999; 99KR-0046795.  
 PR  
 XX (ERUM-) ERUME BIOTECH CO LTD.  
 PA  
 XX Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;  
 XX WPI; 2001-300520/31.  
 DR  
 XX  
 XX New DNA fragments from the rpoB gene of mycobacteria, useful for  
 PT diagnosis and identification of many mycobacterial species by  
 PT restriction fragment length polymorphism -  
 XX  
 XX Claim 1; Page 43; 50pp; English.  
 PS  
 XX The present sequence for Mycobacterium marinum rpoB gene  
 CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from  
 CC various Mycobacterial species. These rpoB gene fragments can be used  
 CC in the diagnosis and identification of Mycobacterium species using a  
 CC novel PCR-restriction fragment length polymorphism analysis (PRA)  
 CC method. The method comprises obtaining a restriction fragment length  
 CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,  
 CC amplifying and digesting the DNA fragment from the microorganism to  
 CC be identified and comparing the RFLP patterns from the known rpoB gene  
 CC fragments with the unidentified fragment. The rpoB gene fragments  
 CC are useful to identify a wide range of Mycobacterium species, e.g. for  
 CC diagnosis or to obtain epidemiological and pathogenesis information for  
 CC selection of appropriate therapies, including M. tuberculosis, M. leprae  
 CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected  
 CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene  
 CC fragments is rapid, precise, simple and cost effective (only 1 PCR  
 CC required), and can differentiate between many species in a single  
 CC experiment, including those difficult to distinguish by usual biochemical  
 CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for  
 CC detecting specific Mycobacterial species.  
 XX  
 SQ Sequence 208 BP; 45 A; 69 C; 68 G; 26 T; 0 other;  
 Query Match 81.5%; Score 169.6; DB 22; Length 208;  
 Best Local Similarity 88.5%; Pred. No. 1e-31; Mismatches 0; Gaps 0;  
 Matches 184; Conservative 0; Mismatches 24; Indels 0; Gaps 0;  
 QY 1 TCAAGGAGAGCGCTACGACCTGGCCGCTGCGCGGATACAAAGTCAACAAGAAGCTGG 60  
 Db 1 TCAAGGAGAGCGCTACGACCTGGCCGCTGCGCGGATACAAAGTCAACAAGAAGCTGG 60  
 QY 61 GCCTGAACACCAATATCTCGATACCAACGACGCTGACCGAAGACGTCGTCGCCA 120  
 Db 61 GCCTGAACCGCGGCCACGCCATCACAGCTCGACGCTGACCGAAGACGTCGTCGCCA 120  
 QY 121 CCATCCAGTATCTGTCGCGCTTCAGAGGCGCCAGCCGATGACCGTCGCGCGGGT 180  
 Db 121 CCATCGAATACCTGGTTCGCTTGCACGAGGCGCCAGACCGGATGACCGTCGCGCGGGT 180  
 QY 181 TCGAGGTGCGGCTGGAAACCGACGACAT 208  
 Db 181 TCGAGGTGCGGCTCGAGACCGACGACAT 208  
 RESULT 6  
 AAS05201

ID AAS05201 standard; DNA; 208 BP.  
 XX  
 AC AAS05201;  
 XX  
 DT 07-SEP-2001 (first entry)  
 XX  
 XX Mycobacterium gordonae type I rpoB gene fragment.  
 DE  
 XX Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;  
 KW PCR-restriction fragment length polymorphism analysis; ds.  
 KW  
 XX Mycobacterium gordonae type I.  
 OS  
 XX WO200131061-A1.  
 PN  
 XX 03-MAY-2001.  
 PD  
 XX 27-OCT-2000; 2000WO-KR01223.  
 PF  
 XX 27-OCT-1999; 99KR-0046795.  
 PR  
 XX (ERUM-) ERUME BIOTECH CO LTD.  
 PA  
 XX Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;  
 PI  
 XX WPI; 2001-300520/31.  
 DR  
 XX  
 XX New DNA fragments from the rpoB gene of mycobacteria, useful for  
 PT diagnosis and identification of many mycobacterial species by  
 PT restriction fragment length polymorphism -  
 XX  
 XX Claim 1; Page 40; 50pp; English.  
 PS  
 XX The present sequence for Mycobacterium gordonae type I rpoB gene  
 CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from  
 CC various Mycobacterial species. These rpoB gene fragments can be used  
 CC in the diagnosis and identification of Mycobacterium species using a  
 CC novel PCR-restriction fragment length polymorphism analysis (PRA)  
 CC method. The method comprises obtaining a restriction fragment length  
 CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,  
 CC amplifying and digesting the DNA fragment from the microorganism to  
 CC be identified and comparing the RFLP patterns from the known rpoB gene  
 CC fragments with the unidentified fragment. The rpoB gene fragments  
 CC are useful to identify a wide range of Mycobacterium species, e.g. for  
 CC diagnosis or to obtain epidemiological and pathogenesis information for  
 CC selection of appropriate therapies, including M. tuberculosis, M. leprae  
 CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected  
 CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene  
 CC fragments is rapid, precise, simple and cost effective (only 1 PCR  
 CC required), and can differentiate between many species in a single  
 CC experiment, including those difficult to distinguish by usual biochemical  
 CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for  
 CC detecting specific Mycobacterial species.  
 XX  
 SQ Sequence 208 BP; 46 A; 71 C; 65 G; 26 T; 0 other;  
 Query Match 80.8%; Score 168; DB 22; Length 208;  
 Best Local Similarity 88.0%; Pred. No. 2.5e-31;  
 Matches 183; Conservative 0; Mismatches 25; Indels 0; Gaps 0;  
 QY 1 TCAAGGAGAGCGCTACGACCTGGCCGCTGTCGCGCCGATACAAAGTCAACAAGAAGCTGG 60  
 Db 1 TCAAGGAGAGCGCTACGACCTGGCCGCTGTCGCGCCGATACAAAGTCAACAAGAAGCTGG 60  
 QY 61 GCCTGAACACCAATATCTCCGATACCAACGACGCTGACCGAAGACGTCGTCGCCA 120  
 Db 61 GCCTGACGTCGCGGATCCGATCACCAGCTCCACGCTGACCGAAGACGTCGTCGCCA 120  
 QY 121 CCATCGAGTATCTGTCGCGCTTCACGAGGCGCCAGCCGATGACCGTCGCGCGGGG 180  
 Db 121 CCATCGAGTATCTGTCGCGCTTCACGAGGCGCCAGCCGATGACCGTCGCGCGGGG 180  
 QY 181 TCGAGGTGCGGTCGAAACCGACGACAT 208

```
Db 181 CCAGGTGCGGTTGAGACCGAGCAT 208
|||||
RESULT 7
AAS05219
ID AAS05219 standard; DNA: 208 BP.
XX
AC AAS05219;
XX
XX 07-SEP-2001 (first entry)
XX
XX Mycobacterium fortuitum rpoB gene fragment.
XX
PN WO200131061-A1.
XX
XX 03-MAY-2001.
XX
XX 27-OCT-2000; 2000WO-KR01223.
XX
XX 27-OCT-1999; 99KR-0046795.
XX
XX (ERUM-) ERUME BIOTECH CO LTD.
XX
XX Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;
XX
XX WPI: 2001-300520/31.
XX
XX New DNA fragments from the rpoB gene of mycobacteria, useful for
PT diagnosis and identification of many mycobacterial species by
PT restriction fragment length polymorphism -
XX
XX Claim 1: Page 46; 50pp; English.
XX
XX The present sequence for Mycobacterium fortuitum rpoB gene
CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
CC various Mycobacterial species. These rpoB gene fragments can be used
CC in the diagnosis and identification of Mycobacterium species using a
CC novel PCR-restriction fragment length polymorphism analysis (PRA)
CC method. The method comprises obtaining a restriction fragment length
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
CC amplifying and digesting the DNA fragment from the microorganism to
CC be identified and comparing the RFLP patterns from the known rpoB gene
CC fragments with the unidentified fragment. The rpoB gene fragments
CC are useful to identify a wide range of Mycobacterium species, e.g. for
CC diagnosis or to obtain epidemiological and pathogenesis information for
CC selection of appropriate therapies, including M. tuberculosis, M. leprae
CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene
CC fragments is rapid, precise, simple and cost effective (only 1 PCR
CC required), and can differentiate between many species in a single
CC experiment, including those difficult to distinguish by usual biochemical
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
CC detecting specific Mycobacterial species.
XX
XX Sequence 208 BP; 43 A; 69 C; 69 G; 27 T; 0 other;
XX
XX Query Match 80.8%; Score 168; DB 22; Length 208;
XX Best Local Similarity 88.0%; Pred. No. 2.5e-31;
XX Matches 183; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

Qy 1 TCAAGGAGAGCGGTAGGACCTGGCGGTGTCGGCGGTACAGGTCACAGAAGCTGG 60
|||||
Db 1 TCAAGGAGAGCGGTAGGACCTGGCGGTGTCGGCGGTACAGGTCACAGAAGCTGG 60
|||||
Qy 61 GCCTGAACCAATCATCCATCACCAGCAGCGTACCAGGAAGAGCGTGTGCGCCA 120
|||||
Db 61 GCCTGAACCGCCGCCGATCAGCTGTCGACTCTGACCCGAGGAAGAGCTGTGCGCCA 120
|||||
```

```
Qy 121 CCATCGAGTATCTGTCGGCTGCACGAGGCCAGGCCAGCATGACCGTCCCGGGGGG 180
|||||
Db 121 CCATCGAGTACCTGGTGGCGCTGCACGAGGCCAGGCCAGCATGACCGTCCCGGGGGG 180
|||||
Qy 181 TCGAGGTGCCGCTGGAACCGACGACAT 208
|||||
Db 181 TCGAGGTCCCGTGCAGGTGACGACAT 208
|||||

RESULT 8
AAS05205
ID AAS05205 standard; DNA: 208 BP.
XX
XX AAS05205;
XX
XX 07-SEP-2001 (first entry)
XX
XX Mycobacterium tuberculosis rpoB gene fragment.
XX
XX Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;
KW PCR-restriction fragment length polymorphism analysis; ds.
KW
XX
XX Mycobacterium tuberculosis.
XX
XX OS WO200131061-A1.
XX
XX PN WO200131061-A1.
XX
XX PD 03-MAY-2001.
XX
XX 27-OCT-2000; 2000WO-KR01223.
XX
XX 27-OCT-1999; 99KR-0046795.
XX
XX (ERUM-) ERUME BIOTECH CO LTD.
XX
XX Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;
XX
XX WPI: 2001-300520/31.
XX
XX New DNA fragments from the rpoB gene of mycobacteria, useful for
PT diagnosis and identification of many mycobacterial species by
PT restriction fragment length polymorphism -
XX
XX Disclosure; Page 41; 50pp; English.
XX
XX The present sequence for Mycobacterium tuberculosis rpoB gene
CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
CC various Mycobacterial species. These rpoB gene fragments can be used
CC in the diagnosis and identification of Mycobacterium species using a
CC novel PCR-restriction fragment length polymorphism analysis (PRA)
CC method. The method comprises obtaining a restriction fragment length
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
CC amplifying and digesting the DNA fragment from the microorganism to
CC be identified and comparing the RFLP patterns from the known rpoB gene
CC fragments with the unidentified fragment. The rpoB gene fragments
CC are useful to identify a wide range of Mycobacterium species, e.g. for
CC diagnosis or to obtain epidemiological and pathogenesis information for
CC selection of appropriate therapies, including M. tuberculosis, M. leprae
CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene
CC fragments is rapid, precise, simple and cost effective (only 1 PCR
CC required), and can differentiate between many species in a single
CC experiment, including those difficult to distinguish by usual biochemical
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
CC detecting specific Mycobacterial species.
XX
XX Sequence 208 BP; 46 A; 62 C; 67 G; 33 T; 0 other;
XX
XX Query Match 80.0%; Score 166.4; DB 22; Length 208;
XX Best Local Similarity 87.5%; Pred. No. 5.9e-31;
XX Matches 182; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

Qy 1 TCAAGGAGAGCGGTAGGACCTGGCGGTGTCGGCGGTACAGGTCACAGAAGCTGG 60
|||||
```

```
Db 1 TCAAGGAGAACGCTACGACCTGGCCCGTCGGTCGTATTAAGGTCAACAAGAGCTCG 60
Qy 61 GCCTGAACACCAATCATCCGATCACCAGGACGCTGACCGAAGAACAGACGTCGTGCGCCA 120
Db 61 GCCTGATGTCGGCGAGCCCATCAGCTGTCGACGCTGACCGAAGAACAGACGTCGTGCGCCA 120
Qy 121 CCATCGAGTATCTGTCGGCTGTCACGAGGCGCCAGGCCACGATGACCGTCGGCGGGGG 180
Db 121 CCATCGAATATCTGTCGGCTGTCACGAGGGTTCAGACACGATGACCGTTCGGCGGGCG 180
Qy 181 TCGAGGTGCGCGTGGAAACCGACGACAT 208
Db 181 TCGAGGTGCGCGTGGAAACCGACGACAT 208

RESULT 9
AAS05206
ID AAS05206 standard; DNA; 208 BP.
XX
AC AAS05206;
XX
DT 07-SEP-2001 (first entry)
XX
DE Mycobacterium terrae rpoB gene fragment.
XX
KW Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;
KW PCR-restriction fragment length polymorphism analysis; ds.
XX
OS Mycobacterium terrae.
XX
PN WO200131061-A1.
XX
PD 03-MAY-2001.
XX
PF 27-OCT-2000; 2000WO-KR01223.
XX
PR 27-OCT-1999; 99KR-0046795.
XX
PA (ERUM-) ERUME BIOTECH CO LTD.
XX
PI Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;
XX
PS WPI; 2001-300520/31.
XX
CC New DNA fragments from the rpoB gene of mycobacteria, useful for
CC diagnosis and identification of many mycobacterial species by
CC restriction fragment length polymorphism.
CC
CC Claim 1; Page 42; 50pp; English.
CC
CC The present sequence for Mycobacterium terrae rpoB gene
CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
CC various Mycobacterial species. These rpoB gene fragments can be used
CC in the diagnosis and identification of Mycobacterium species using a
CC novel PCR-restriction fragment length polymorphism analysis (PRA)
CC method. The method comprises obtaining a restriction fragment length
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
CC amplifying and digesting the DNA fragment from the microorganism to
CC be identified and comparing the RFLP patterns from the known rpoB gene
CC fragments with the unidentified fragment. The rpoB gene fragments
CC are useful to identify a wide range of Mycobacterium species, e.g. for
CC diagnosis or to obtain epidemiological and pathogenesis information for
CC selection of appropriate therapies, including M. tuberculosis, M. leprae
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene
CC fragments is rapid, precise, simple and cost effective (only 1 PCR
CC required), and can differentiate between many species in a single
CC experiment, including those difficult to distinguish by usual biochemical
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
CC detecting specific Mycobacterial species.
XX
SQ Sequence 208 BP; 46 A; 62 C; 67 G; 33 T; 0 other;
```

```
Query Match 80.0%; Score 166.4; DB 22; Length 208;
Best Local Similarity 87.5%; Pred. No. 5.9e-31;
Matches 182; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

Qy 1 TCRAGGAGAACGCTACGACCTGGCCCGTCGGTCGTATTAAGGTCAACAAGAGCTCG 60
Db 1 TCRAGGAGAACGCTACGACCTGGCCCGTCGGTCGTATTAAGGTCAACAAGAGCTCG 60
Qy 61 GCCTGAACACCAATCATCCGATCACCAGGACGCTGACCGAAGAACAGACGTCGTGCGCCA 120
Db 61 GCCTGATGTCGGCGAGCCCATCAGCTGTCGACGCTGACCGAAGAACAGACGTCGTGCGCCA 120
Qy 121 CCATCGAGTATCTGTCGGCTGTCACGAGGCGCCAGGCCACGATGACCGTCGGCGGGGG 180
Db 121 CCATCGAATATCTGTCGGCTGTCACGAGGGTTCAGACACGATGACCGTTCGGCGGGCG 180
Qy 181 TCGAGGTGCGCGTGGAAACCGACGACAT 208
Db 181 TCGAGGTGCGCGTGGAAACCGACGACAT 208

RESULT 10
AAS05210
ID AAS05210 standard; DNA; 208 BP.
XX
AC AAS05210;
XX
DT 07-SEP-2001 (first entry)
XX
DE Mycobacterium ulcerans rpoB gene fragment.
XX
KW Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;
KW PCR-restriction fragment length polymorphism analysis; ds.
XX
OS Mycobacterium ulcerans.
XX
PN WO200131061-A1.
XX
PD 03-MAY-2001.
XX
PF 27-OCT-2000; 2000WO-KR01223.
XX
PR 27-OCT-1999; 99KR-0046795.
XX
PA (ERUM-) ERUME BIOTECH CO LTD.
XX
PI Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;
XX
PS WPI; 2001-300520/31.
XX
CC New DNA fragments from the rpoB gene of mycobacteria, useful for
CC diagnosis and identification of many mycobacterial species by
CC restriction fragment length polymorphism.
CC
CC Claim 1; Page 43; 50pp; English.
CC
CC The present sequence for Mycobacterium ulcerans rpoB gene
CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
CC various Mycobacterial species. These rpoB gene fragments can be used
CC in the diagnosis and identification of Mycobacterium species using a
CC novel PCR-restriction fragment length polymorphism analysis (PRA)
CC method. The method comprises obtaining a restriction fragment length
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
CC amplifying and digesting the DNA fragment from the microorganism to
CC be identified and comparing the RFLP patterns from the known rpoB gene
CC fragments with the unidentified fragment. The rpoB gene fragments
CC are useful to identify a wide range of Mycobacterium species, e.g. for
CC diagnosis or to obtain epidemiological and pathogenesis information for
CC selection of appropriate therapies, including M. tuberculosis, M. leprae
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene
CC fragments is rapid, precise, simple and cost effective (only 1 PCR
CC required), and can differentiate between many species in a single
CC experiment, including those difficult to distinguish by usual biochemical
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
CC detecting specific Mycobacterial species.
XX
SQ Sequence 208 BP; 46 A; 62 C; 67 G; 33 T; 0 other;
```

CC required), and can differentiate between many species in a single  
CC experiment, including those difficult to distinguish by usual biochemical  
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for  
CC detecting specific Mycobacterial species.

XX  
SQ Sequence 208 BP; 45 A; 68 C; 67 G; 28 T; 0 other;  
Query Match 80.0%; Score 166.4; DB 22; Length 208;  
Best Local Similarity 87.5%; Pred. No. 5.9e-31;  
Matches 182; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 1 TCAAGGAGAGCGCTACGACCTGCGCCGCTGTCGCGCGATACAAGGTCAACAAGAGCTGG 60  
DB 1 TCAAGGAGAGCGCTACGACCTGCGCTGCGTACAAAGGTCAACAAGAGCTCG 60  
QY 61 GCCTGACACCAATCATCCGATCACCAGCAGCAGCTGACCGAAGAAGAGCTGTCGCCA 120  
DB 61 GCCTGAAAGCGCGCCAGCCCATCACCAGCTCGACGCTGACCGAGGAAGAGCTGTCGCCA 120  
QY 121 CCATCGAGTATCTGCTCGGCTGCACGAGGCGCCAGGCGACGATGACCGTCCCGGGCGGG 180  
DB 121 CCATCGAATACCTGCTCGGCTGTCACGAGGCGCCAGACCGCGATGACCGCTCCGGCGGTG 180  
QY 181 TCGAGGTCCCGGTGGAACCGAGACAT 208  
DB 181 TCGAGGTCCCGTCCGAGACCGAGACAT 208

## RESULT 11

AAS05216  
ID AAS05216 standard; DNA; 208 BP.

XX AC AAS05216;

XX DT 07-SEP-2001 (first entry)

XX DE Mycobacterium bovis rpoB gene fragment.

XX KW Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;  
XX PCR-restriction fragment length polymorphism analysis; ds.

XX OS Mycobacterium bovis.

XX PN WO200131061-A1.

XX PD 03-MAY-2001.

XX PF 27-OCT-2000; 2000WO-KR01223.

XX PR 27-OCT-1999; 99KR-0046795.

XX PA (ERUM-) ERUME BIOTECH CO LTD.

XX PI Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;

XX DR WPI; 2001-300520/31.

XX PT New DNA fragments from the rpoB gene of mycobacteria, useful for  
XX diagnosis and identification of many mycobacterial species by  
XX restriction fragment length polymorphism -

XX PS Claim 1; Page 45; 50pp; English.

XX CC The present sequence for Mycobacterium bovis rpoB gene  
CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from  
CC various Mycobacterial species. These rpoB gene fragments can be used  
CC in the diagnosis and identification of Mycobacterium species using a  
CC novel PCR-restriction fragment length polymorphism analysis (PRA)  
CC method. The method comprises obtaining a restriction fragment length  
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,  
CC amplifying and digesting the DNA fragment from the microorganism to  
CC be identified and comparing the RFLP patterns from the known rpoB gene  
CC fragments with the unidentified fragment. The rpoB gene fragments

CC are useful to identify a wide range of Mycobacterium species, e.g. for  
CC diagnosis or to obtain epidemiological and pathogenesis information for  
CC selection of appropriate therapies, including M. tuberculosis, M. leprae  
CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected  
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene  
CC fragments is rapid, precise, simple and cost effective (only 1 PCR  
CC experiment), and can differentiate between many species in a single  
CC experiment, including those difficult to distinguish by usual biochemical  
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for  
CC detecting specific Mycobacterial species.

XX SQ Sequence 208 BP; 46 A; 62 C; 67 G; 33 T; 0 other;

Query Match 80.0%; Score 166.4; DB 22; Length 208;  
Best Local Similarity 87.5%; Pred. No. 5.9e-31;  
Matches 182; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 1 TCAAGGAGAGCGCTACGACCTGCGCCGCTGTCGCGCGATACAAGGTCAACAAGAGCTGG 60  
DB 1 TCAAGGAGAGCGCTACGACCTGCGCCGCTGCGTATTAAGGTCAACAAGAGCTCG 60

QY 61 GCCTGACACCAATCATCCGATCACCAGCAGCAGCTGACCGAAGAAGAGCTGTCGCCA 120  
DB 61 GCCTGATGTCGCGAGCCCATCAGCTGCTGACGCTGACCGAAGAAGAGCTGTCGCCA 120

QY 121 CCATCGAGTATCTGCTCGCCCTGCACGAGGCGCCAGCGATGACCGTCCCGGGCGGG 180  
DB 121 CCATCGAATATCTGCTCGCCTTGACAGAGGTTCACACCGATGACCGTCCCGGGCGGG 180

QY 181 TCGAGGTCCCGGTGGAACCGAGACAT 208

DB 181 TCGAGGTCCCGTGGAAACCGAGACAT 208

## RESULT 12

AAT09676  
ID AAT09676 standard; DNA; 970 BP.

XX AC AAT09676;

XX DT 15-OCT-1996 (first entry)

XX DE Mycobacterium tuberculosis rpoB gene DNA sequence.

XX KW Tuberculosis; disease diagnosis; oligonucleotide; DNA primer; PCR;  
XX polymerase chain reaction; DNA amplification; rpoB locus; TB; sa.

XX OS Mycobacterium tuberculosis.

XX FH Key Location/Qualifiers  
XX primer\_bind 10..27  
FT /tag= a  
FT /note= "primer FENLFF"

FT primer\_bind 226..243  
FT /tag= b  
FT /note= "primer DDIDLH"

FT primer\_bind 226..240  
FT /tag= c  
FT /note= "primer DDIDLH"

FT primer\_bind 338..364  
FT /tag= d  
FT /note= "primer rpo95"

FT primer\_bind 348..373  
FT /tag= e  
FT /note= "primer rpo105"

FT primer\_bind 354..373  
FT /tag= f  
FT /note= "primer KY290"

FT misc\_feature 372..373  
FT /tag= g  
FT /note= "M. tuberculosis signature nucleotide"

FT misc\_feature 433..434  
FT /tag= h



```
SQ Sequence 3519 BP; 675 A; 1081 C; 1183 G; 580 T; 0 other;
Query Match 80.0%; Score 166.4; DB 22; Length 3519;
Best Local Similarity 87.5%; Pred. No. 7.4e-31;
Matches 182; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

Oy 1 TCAAGGAGAGCGCTAGACCTGCGCGTGTGCGCGGATACAAAGGTCAACAAGAGCTGG 60
Db 884 TCAAGGAGAGCGCTAGACCTGCGCGTGTGCGCGGATACAAAGGTCAACAAGAGCTGG 943
Oy 61 GCCTGACACCAATCATCCGATCACCACGAGGAGCTGACCGAAGAGAGCTGTCGCCCA 120
Db 944 GCCTGACATGTGCGGAGCCCATCAGCTCGTCGACGCTGACCGAAGAGAGCTGTCGCCCA 1003
Oy 121 CCATCGAGTATCTGCTCGGCTGCACGAGGCGGACGACGATGACCGTCCCGGGGGG 180
Db 1004 CCATCGAATATCTGCTCGGCTGCACGAGGTCAGACCAACGATGACCGTCCCGGGGGG 1063
Oy 181 TCGAGGTGCGGTGGAAACCGACGACAT 208
Db 1084 TCGAGGTGCGGTGGAAACCGACGACAT 1091

RESULT 14
ID AAH02079 standard; DNA: 3534 BP.
XX AC AAH02079;
XX DT 24-JUL-2001 (first entry)
XX DE Mycobacterium tuberculosis nucleotide sequence SEQ ID NO:2072.
XX KW Species specific; genus specific; family specific; probe: detection;
XX KW identification; algal; archaeal; bacterial; fungal; parasitica;
XX KW microorganism; diagnosis; translation elongation factor Tu; toxin;
XX KW translation elongation factor G; RecA recombinase; resistance;
XX KW catalytic subunit of proton-translocating ATPase; antimicrobial;
XX KW vaccine; primer; ds.
XX OS Mycobacterium tuberculosis.
XX PN WO200123604-A2.
XX PD 05-APR-2001.
XX PF 28-SEP-2000; 2000WO-CA01150.
XX PR 28-SEP-1999; 99CA-2283458.
XX PR 19-MAY-2000; 2000CA-2307010.
XX PA (INFE-) INFECTIO DIAGNOSTIC (IDI) INC.
XX PI Bergeron MG, Boissinot M, Huletsky A, Menard C, Ouellette M;
XX PI Picard FJ, Roy PH;
XX DR WPI: 2001-245006/25.
XX PT Nucleic acid sequences are used to generate universal probes and
XX PT primers which can be used to identify and detect the presence of algal,
XX PT archaeal, bacterial, fungal and parasitica species in a test sample -
XX PS Disclosure; Page 1478-1479; 1580pp; English.
XX CC The present invention describes a method for generating a repertory of
XX CC nucleic acids of tuf, fus, atpD and/or recA genes from which probes
XX CC and/or primers are derived. The method comprises amplifying the nucleic
XX CC acids of determined algal, archaeal, bacterial, fungal and parasitica
XX CC species with a combination of defined primer pairs. The method can be
XX CC used for producing probes and/or primers for detecting one or more
XX CC related microorganisms e.g. algae, archaea, bacteria, fungi and
XX CC parasites, for universal detection and for specific and ubiquitous
XX CC detection and identification of an algal, archaeal, bacterial, fungal
```

---

```
CC and parasitica species, genus, family and group. A nucleic acid (1)
CC obtained using the method of the invention can be used for the universal
CC detection of any bacterium, fungus or parasite in a sample and for the
CC detection of at least one antimicrobial agent resistance gene or at
CC least one toxin gene. hexA nucleic acids are used for the specific and
CC ubiquitous detection and for identification of Streptococcus pneumoniae.
CC (1) can be used to design a therapeutic agent which is effective against
CC microorganisms. Microbial species or genus or family or phylum or group
CC which can be detected include Abiotrophia adiacens, Bordetella sp.,
CC Corynebacterium sp., Enterobacteriaceae group, Escherichia coli,
CC Mycobacteriaceae family, Pseudomonas group, Streptococcus sp.,
CC Neisseria gonorrhoeae and Staphylococcus sp.. Using DNA based tests
CC provides faster results than substrate specificity tests as results can
CC be determined in an hour and improved accuracy is also achieved.
CC AAH00010 to AAH002304 represent nucleotide sequences and primers/probes
CC which are given in the exemplification of the present invention.
XX Sequence 3534 BP; 679 A; 1081 C; 1188 G; 586 T; 0 other;
Query Match 80.0%; Score 166.4; DB 22; Length 3534;
Best Local Similarity 87.5%; Pred. No. 7.4e-31;
Matches 182; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

Oy 1 TCAAGGAGAGCGCTAGACCTGCGCGTGTGCGCGGATACAAAGGTCAACAAGAGCTGG 60
Db 902 TCAAGGAGAGCGCTAGACCTGCGCGTGTGCGCGGATACAAAGGTCAACAAGAGCTGG 961
Oy 61 GCCTGACACCAATCATCCGATCACCACGAGGAGCTGACCGAAGAGAGCTGTCGCCCA 120
Db 962 GCCTGATGTGCGGAGCGCCATCAGCTGCTGCGAGCTGACCGAAGAGAGCTGTCGCCCA 1021
Oy 121 CCATCGAGTATCTGCTCGGCTGCACGAGGCGGACCGACGATGACCGTCCCGGGGGG 180
Db 1022 CCATCGAATATCTGCTCGGCTGCACGAGGTCAGACCAACGATGACCGTCCCGGGGGG 1081
Oy 181 TCGAGGTGCGGTGGAAACCGACGACAT 208
Db 1082 TCGAGGTGCGGTGGAAACCGACGACAT 1109

RESULT 15
ID AAA74651 standard; DNA: 3853 BP.
XX AC AAA74651;
XX DT 06-DEC-2000 (first entry)
XX DE Mycobacterium tuberculosis rpoB gene.
XX KW Mycobacterium tuberculosis; rpoB; RNA polymerase beta subunit;
XX KW rifampin resistance; mutation detection; ds.
XX OS Mycobacterium tuberculosis.
XX PN WO200043546-A2.
XX PD 27-JUL-2000.
XX PF 20-DEC-1999; 99WO-US30377.
XX PR 19-JAN-1999; 99US-0233996.
XX PA (DADE-) DADE BEHRING INC.
XX PI Liu YP, Kurn N;
XX DR WPI: 2000-524243/47.
XX PT Method for detecting drug resistance in a strain of an organism,
XX PT particularly for detecting rifampin resistance in Mycobacterium
XX PT tuberculosis -
```



PS Example 1; Fig 4; 86pp; English.

xx  
CC The present sequence is the rpoB gene from Mycobacterium tuberculosis.  
CC Rifampin resistance is largely associated with point mutations  
CC localised in a small core region of 81 base pairs in the rpoB gene, which  
CC encodes the RNA polymerase beta subunit. To detect a mutation, a complex  
CC is formed comprising a first sequence representing the predetermined  
CC region of the gene of the organism and a second sequence representing the  
CC corresponding region of the gene of the wild type organism in double  
CC stranded form. Each member of at least one pair of non-complementary  
CC strands within the complex has a label. The association of the labels in  
CC the complex is related to the presence of the mutation. The presence of  
CC the mutation is related to the drug resistance of the strain.

xx SQ Sequence 3853 BP; 723 A; 1173 C; 1293 G; 664 T; 0 other;

Query Match 80.0%; Score 166.4; DB 21; Length 3853;  
Best Local Similarity 87.5%; Pred. No. 7.5e-31;  
Matches 182; Conservative 0; Mismatches 26; Indels 0; Gaps 0;  
Oy 1 TCAAGGAGAACGCTACGACCTGGCCGCTGTGCGCCGATACAGGTCAACAAGAAGCTGG 60  
Db 1477 TCAAGGAGAACGCTACGACCTGGCCGCTGTGCGCTCGCTATAGGTCAACAAGAAGCTCG 1536  
Oy 61 GCCTGAACACCAATCATCCGATCACACGACGCTGACCGAAGAGAGAGCTCGTCGCCA 120  
Db 1537 GGCTGCATGTCCGGCAGCCCATACGTCGTGACGCTGACCGAAGAGAGCTCGTCGCCA 1596  
Oy 121 CCATCGAGTATCTGGTCGCCCTGCACGAGGCCAGGCCAGTACCGTCCGGGGGGGG 180  
Db 1597 CCATCGAATATCTGGTCGCCCTTGCACGAGGGTCAGACGATGACCGTTCCGGGGGGCG 1656  
Oy 181 TCGAGGTGCGGTGGAACCGACGACAT 208  
Db 1657 TCGAGGTGCGGTGGAACCGACGACAT 1684

Search completed: November 12, 2002, 16:50:09  
Job time : 141.222 secs